

**«Μελέτη της διαφορικής έκφρασης
γονιδίων έπειτα από έκθεση σε
χαμηλές δόσεις ιοντιζουσών
ακτινοβολιών μέσω ανάλυσης
δεδομένων μικροσυστοιχιών»**



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ
ΠΟΛΥΤΕΧΝΕΙΟ**

ΘΕΟΔΩΡΑ-ΔΑΦΝΗ ΜΙΧΑΛΕΤΤΟΥ

Επιβλέπων:

Αλέξανδρος Γεωργακίλας
Αναπληρωτής Καθηγητής Ε.Μ.Π.

**ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ
ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ**

Διπλωματική Εργασία
Αθήνα 2017



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

**«Μελέτη της διαφορικής έκφρασης γονιδίων έπειτα από
έκθεση σε χαμηλές δόσεις ιοντιζουσών ακτινοβολιών μέσω
ανάλυσης δεδομένων μικροσυστοιχιών»**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΘΕΟΔΩΡΑ-ΔΑΦΝΗ ΜΙΧΑΛΕΤΤΟΥ

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

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

Τριμελής εξεταστική επιτροπή:

Αλέξανδρος Γεωργακίλας
Αναπληρωτής Καθηγητής
Ε.Μ.Π.

(Υπογραφή)

.....

Κων/νος Βοργιάς
Καθηγητής Βιοχημείας
Ε.Κ.Π.Α.

(Υπογραφή)

.....

Παντελής Μπάγκος
Αναπληρωτής Καθηγητής
Παν. Θεσσαλίας

(Υπογραφή)

.....

Αθήνα 2017

Ευχαριστίες

Η βάση δεδομένων «microarray» υλοποιήθηκε στο Ίδρυμα Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών (ΙΙΒΕΑΑ). Ευχαριστώ τον Αναπ. Καθ. κ. Α. Γεωργακίλα για την ανάθεση του θέματος και την καθοδήγηση του κατά την διεξαγωγή της εργασίας αυτής. Ευχαριστώ θερμά τον Δρ. Ιωάννη Μιχαλόπουλο του ΙΙΒΕΑΑ, με την επίβλεψη και διαρκή καθοδήγηση του οποίου ολοκληρώθηκε η διπλωματική μου εργασία ενώ παράλληλα συνέβαλε στο να κατανοήσω την Βιοπληροφορική προσέγγιση του θέματος που πραγματεύτηκα. Ευχαριστώ επίσης τον Καθ. κ. Π. Μπάγκο και την ερευνητική του ομάδα για την συνεργασία και καθοδήγηση σε θέματα μετα-ανάλυσης. Στη συνέχεια ευχαριστώ εκ νέου τον Αναπ. Καθ. κ. Π. Μπάγκο αλλά και τον Καθ. κ. Κ. Βοργιά, που δέχθηκαν να συμμετέχουν στην τριμελή εξεταστική επιτροπή μου. Τέλος, θα ήθελα να ευχαριστήσω τους γονείς μου για την αμέριστη συμπαράσταση τους κατά την διάρκεια των σπουδών μου, τον αδερφό μου για τις εποικοδομητικές συζητήσεις στα θέματα βιοχημείας, τον φίλο μου Γιάννη για την συναισθηματική υποστήριξη αλλά και την σύντροφο Χριστίνα για την πάσης φύσεως βοήθεια.

Περίληψη

Εισαγωγή

Η χρήση ιοντιζουσών ακτινοβολιών αποτελεί την συνηθέστερη μέθοδο προσέγγισης για την διάγνωση και την θεραπεία πολλών ασθενειών, συμπεριλαμβανομένων των περισσότερων τύπων καρκίνου. Σύμφωνα με πολλαπλές πηγές, θεωρείται ότι η επίδραση των υψηλών δόσεων ακτινοβολίας σχετίζεται κυρίως με την φλεγμονή και το οξειδωτικό στρες. Στόχος της εργασίας είναι ο καθορισμός των επιδράσεων των χαμηλών δόσεων ακτινοβολίας στον ανθρώπινο ιστό. Για το σκοπό αυτό, συγκρίνονται τα διαφορικά εκφραζόμενα γονίδια ως απόκριση στην έκθεση σε χαμηλές και υψηλές δόσεις καθώς και τα βιολογικά δίκτυα στα οποία συμμετέχουν αυτά.

Μέθοδος

Επιλέχθηκαν δείγματα από ανθρώπινους ιστούς κατηγοριοποιημένα ως ακτινοβολημένα με χαμηλή ή υψηλή δόση ακτινοβολίας και μη ακτινοβολημένα (δείγματα ελέγχου) και καθορίστηκε σε αυτά η γονιδιακή έκφραση μέσω ανάλυσης δεδομένων μικροσυστοιχιών. Τα πρωταρχικά δεδομένα από τις μελέτες λήφθηκαν από το GEO (δημόσιο αποθετήριο γονιδιακής έκφρασης). Τα δεδομένα (από τις 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 που βασίζεται στην γλώσσα προγραμματισμού R. Επίσης λήφθηκαν γονιδιακά δεδομένα από το BioMart και σχεδιάστηκε βάση δεδομένων MySQL με σκοπό την αποθήκευση των δεδομένων μας και την διεξαγωγή μετα-ανάλυσης για την επιλογή των στατιστικά σημαντικών διαφορικά εκφρασμένων γονιδίων. Χρησιμοποιώντας την STRING και το WebGestalt, καταφέραμε να σχεδιάσουμε γονιδιακά δίκτυα και να καθορίσουμε τις βιολογικές διεργασίες που σχετίζονται με την απόκριση στις ιοντιζουσες ακτινοβολίες.

Αποτελέσματα

Στα αποτελέσματα της ανάλυσης αυτής για την επίδραση των χαμηλών δόσεων ιοντιζουσας ακτινοβολίας συγκαταλέγονται: η ενεργοποίηση μηχανισμών επιδιόρθωσης DNA και διαδικασίες που σχετίζονται με την αναστολή του κυτταρικού κύκλου καθώς και η ενεργοποίηση απόκρισης έπειτα από έκθεση σε ακτινοβολία ιοντιζουσα και μη. Μεγαλύτερο ενδιαφέρον όμως παρουσιάζει η ενεργοποίηση του μονοπατιού «GAB1 *Signalosome*», που οδηγεί στην ενεργοποίηση μέσω σηματοδότησης της AKT, η οποία και έχει συσχετιστεί με την δημιουργία καρκίνου.

Τέλος, υπάρχουν ενδείξεις για μονοπάτια που μπορούν να συσχετιστούν με την επεξήγηση του φαινομένου του παρατυχόντος.

Συζήτηση

Συγκρίνοντας την απόκριση σε υψηλές και χαμηλές δόσεις ακτινοβολίας γίνεται αντιληπτό ότι υπάρχουν διαφορές. Ενώ στις υψηλές είναι βέβαιη η κινητοποίηση του κυττάρου μέσω της ενεργοποίησης της p53 και ορισμένων μηχανισμών επιδιόρθωσης DNA, στις χαμηλές δόσεις η απόπειρα του κυττάρου να ενεργήσει κατάλληλα δεν διαφαίνεται πλήρως. Παρόλα αυτά η εμφάνιση του γονιδίου CDKN1A στην πλειοψηφία των αποτελεσμάτων για την έκθεση σε χαμηλές δόσεις, η έκφραση του οποίου οδηγεί στην παραγωγή της p53-εξαρτώμενης πρωτεΐνης p21, δίνει ενδείξεις για την ενεργοποίηση της ογκο-κατασταλτικής πρωτεΐνης p53 και των αντίστοιχων μηχανισμών επιδιόρθωσης και στην περίπτωση των χαμηλών δόσεων.

Αν και τα γονίδια που προέκυψαν από την ανάλυση ως στατιστικώς σημαντικά διαφορετικώς εκφρασμένα μπορούν να συσχετιστούν με την βιβλιογραφία, χρήζουν αναθεώρησης καθώς δεν αναδείχθηκαν από τις μεθόδους διόρθωσης πολλαπλών ερωτημάτων. Η στατιστική αξία αυτών θα ήταν μεγαλύτερη αν είχαν προέλθει από μεγαλύτερο αριθμό σειρών δεδομένων στην περίπτωση του μοντέλου τυχαίων επιδράσεων κατά την διεξαγωγή μετα-ανάλυσης, ή αν ο σχεδιασμός των πειραμάτων χαρακτηριζόταν από μεγαλύτερη ομοιογένεια, επιλέγοντας έτσι το μοντέλο των σταθερών επιδράσεων. Επίσης η χρήση μελετών που χρησιμοποιούσαν κυτταρικές σειρές πιθανά επηρέασε την ανάλυση, καθώς στα αθανатоποιημένα κύτταρα έχουν επηρεαστεί οι διεργασίες του κυτταρικού κύκλου, οι οποίες αποτελούν βασικό τμήμα της απόκρισης σε ιοντίζουσα ακτινοβολία. Τέλος οφείλει να σημειωθεί ότι η διεξαγωγή επιπλέον πειραμάτων για τον καθορισμό των επιδράσεων των χαμηλών δόσεων ιοντίζουσας ακτινοβολίας είναι αναγκαία για την βελτίωση των μεθόδων διάγνωσης και θεραπείας και την πρόληψη πολλών ασθενειών.

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

Abstract

Introduction

The use of Ionizing Radiation (IR) is currently one of the most common approaches in the diagnosis and treatment of many kinds of disease, including most types of cancer. Evidence from multiple sources suggest the involvement of inflammation and oxidative stress to be the main contributors of high dose radiation effects. Our interest was to identify the effects of low dose radiation in human tissues. To this end, we compare up- or down-regulated genes in response to low and high dose and the biological networks they participate in.

Materials and Methods

Human samples categorized as irradiated with low or high dose of radiation and non-irradiated (control samples) were selected and gene expression was determined using microarray data analysis. Raw data from various studies were downloaded from GEO public microarray database (GEO series GSE8917; GSE12435; GSE16935; GSE23901; GSE29344; GSE52918; GSE59861). GEO data were background-corrected, quantile normalized and \log_2 transformed, using the Limma package from Bioconductor suite which is based on the R programming language. We also downloaded various genomic data from BioMart. We designed a MySQL database in order to store our data and conduct a meta-analysis to select statistically significant differentially expressed genes. Using STRING and WebGestalt, we were able to construct gene networks and identify biological processes which are related to IR response.

Results

The results of this analysis for the effect of low doses of ionizing radiation include: activation of DNA repair mechanisms and processes related to cell cycle arrest as well as the response after exposure to ionising and non ionising radiation. Also the activation of the GAB1 Signaling pathway, leads to AKT signaling that has been associated with cancer. Finally, there are indications of pathways that can be related to the explanation of the bystander effect.

Discussion

The differences in response to high and low doses of radiation are evident. While at high doses the activation of p53 and several DNA repair mechanisms ensure the cell's response to radiation, the corresponding response of the cell to low doses is not entirely clear. However, the appearance of CDKN1A gene in the majority of results for exposure to low dose radiation, the expression of which leads to the production of p53-dependent p21 protein, gives evidence of activation of the p53 tumor suppressor protein and all the corresponding repair mechanisms in the case of low doses respectively.

Although the genes that stand out from this analysis as statistically significant differentially expressed may be correlated with literature, they certainly need revision because they failed the methods of multiple tests adjustment. The statistical value of these genes would be greater if they derived from a larger number of data sets in the case of the random effect model of meta-analysis, or if the entirety of the selected experiments was characterized by greater homogeneity, in this case choosing the fixed effect model. In addition, choosing studies that used samples from cell line cultures may have affected this analysis because cell cycle is a key part of ionizing radiation response and in immortalized cells the cell cycle process has been disturbed. Finally, it is very important to proceed to further experiments in order to determine the effects of low dose ionizing radiation, as they are necessary for improving diagnose and treatment methods and also prevent many kinds of disease.

Keywords: ionizing radiation, low dose, high dose, gene differential expression, bystander effect, bioinformatics, microarray analysis, database, meta-analysis, biomarkers, biological networks

ΕΙΚΟΝΑ 1: ΠΑΡΑΓΩΓΗ ΑΚΤΙΝΑΣ-Γ ΜΕΣΩ ΑΠΟΔΙΕΓΕΡΣΗΣ ΔΙΕΓΕΡΜΕΝΗΣ ΚΑΤΑΣΤΑΣΗΣ ΠΛΟΥΤΩΝΙΟΥ (PU), ΟΠΟΥ Ο ΜΗΤΡΙΚΟΣ ΠΥΡΗΝΑΣ ΔΕΝ ΑΛΛΑΖΕΙ ΩΣ ΠΡΟΣ ΤΗΝ ΜΑΖΑ ΚΑΙ ΤΟΝ ΑΤΟΜΙΚΟ ΑΡΙΘΜΟ.	4
ΕΙΚΟΝΑ 2: ΜΕΤΑΠΗΔΗΣΗ ΔΙΕΓΕΡΜΕΝΟΥ ΗΛΕΚΤΡΟΝΙΟΥ ΑΠΟ ΤΙΣ ΣΤΟΙΒΑΔΕΣ L ΚΑΙ M ΣΤΗΝ Κ ΠΟΥ ΣΥΝΤΕΛΕΙ ΣΤΗΝ ΕΚΠΟΜΠΗ ΑΚΤΙΝΩΝ-X.....	4
ΕΙΚΟΝΑ 3: Α-ΔΙΑΣΠΑΣΗ ΟΥΡΑΝΙΟΥ ΠΟΥ ΣΥΝΕΠΑΓΕΤΑΙ ΠΑΡΑΓΩΓΗ ΣΩΜΑΤΙΔΙΟΥ-Α ΚΑΙ ΘΥΓΑΤΡΙΚΟΥ ΠΥΡΗΝΑ ΡΑΔΙΕΝΕΡΓΟΥ ΘΟΡΙΟΥ.....	5
ΕΙΚΟΝΑ 4: Β-ΔΙΑΣΠΑΣΗ ΘΟΡΙΟΥ ΠΟΥ ΣΥΝΕΠΑΓΕΤΑΙ ΠΑΡΑΓΩΓΗ ΣΩΜΑΤΙΔΙΟΥ-Β (ΗΛΕΚΤΡΟΝΙΟΥ) ΚΑΙ ΔΙΕΓΕΡΜΕΝΟΥ ΘΥΓΑΤΡΙΚΟΥ ΠΥΡΗΝΑ ΠΟΥ ΑΠΟΔΙΕΓΕΙΡΕΤΑΙ ΜΕ ΕΚΠΟΜΠΗ ΦΩΤΟΝΙΟΥ.	6
ΕΙΚΟΝΑ 5: (Α) ΤΡΟΧΙΑ ΙΟΝΤΙΣΜΩΝ LOW LET ΑΚΤΙΝΟΒΟΛΙΑΣ ΚΑΙ (Β) ΤΡΟΧΙΑ ΙΟΝΤΙΣΜΩΝ HIGH LET ΑΚΤΙΝΟΒΟΛΙΑΣ ΣΕ ΕΝΑ ΚΥΤΤΑΡΟ.....	9
ΕΙΚΟΝΑ 6: ΑΝΤΙΔΡΑΣΕΙΣ ΠΑΡΑΓΩΓΗΣ ΕΛΕΥΘΕΡΩΝ ΡΙΖΩΝ.....	11
ΕΙΚΟΝΑ 7: ΑΝΤΙΔΡΑΣΕΙΣ ΡΑΔΙΟΛΥΣΗΣ ΤΟΥ ΥΔΑΤΟΣ.	11
ΕΙΚΟΝΑ 8: ΔΙΑΧΩΡΙΣΜΟΣ ΣΤΟΧΑΣΤΙΚΩΝ ΚΑΙ ΜΗ-ΣΤΟΧΑΣΤΙΚΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΙΟΝΤΙΖΟΥΣΑ ΑΚΤΙΝΟΒΟΛΙΑ.	12
ΕΙΚΟΝΑ 9: ΔΙΑΧΩΡΙΣΜΟΣ ΣΤΟΧΕΥΜΕΝΩΝ ΚΑΙ ΜΗ ΣΤΟΧΕΥΜΕΝΩΝ ΕΠΙΔΡΑΣΕΩΝ ΤΗΣ ΙΟΝΤΙΖΟΥΣΑΣ ΑΚΤΙΝΟΒΟΛΙΑΣ.	13
ΕΙΚΟΝΑ 10: ΔΙΑΓΡΑΜΜΑ ΑΝΑΠΑΡΑΣΤΑΣΗΣ ΤΟΥ ΚΥΤΤΑΡΙΚΟΥ ΚΥΚΛΟΥ.	14
ΕΙΚΟΝΑ 11: ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΔΙΑΦΟΡΩΝ ΠΙΘΑΝΩΝ ΣΕΝΑΡΙΩΝ ΓΙΑ ΤΟΥΣ ΚΙΝΔΥΝΟΥΣ ΑΚΤΙΝΟΒΟΛΙΑΣ ΕΠΕΙΤΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΠΟΛΥ ΧΑΜΗΛΕΣ ΔΟΣΕΙΣ, ΟΙ ΟΠΟΙΕΣ ΜΠΟΡΕΙ ΝΑ ΕΙΝΑΙ ΣΥΜΒΑΤΕΣ ΜΕ ΕΠΙΔΗΜΙΟΛΟΓΙΚΑ ΔΕΔΟΜΕΝΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΟΤΕΡΕΣ ΔΟΣΕΙΣ ΑΚΤΙΝΟΒΟΛΙΑΣ ΌΛΕΣ ΟΙ ΣΧΕΣΕΙΣ ΔΟΣΗΣ-ΑΠΟΚΡΙΣΗΣ ΑΠΟΤΕΛΟΥΝ ΠΙΘΑΝΕΣ ΠΕΡΙΓΡΑΦΕΣ ΟΓΚΟΓΕΝΕΣΗΣ (BRENNER ET AL., 2003).....	19
ΕΙΚΟΝΑ 12: ΕΞΙΔΑΝΙΚΕΥΜΕΝΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΔΙΑΦΟΡΩΝ ΤΥΠΩΝ ΟΜΑΔΟΠΟΙΗΜΕΝΩΝ ΒΛΑΒΩΝ, ΣΥΜΠΕΡΙΛΑΜΒΑΝΟΜΕΝΩΝ ΣΥΝΘΕΤΩΝ ΘΡΑΥΣΜΑΤΩΝ ΜΟΝΗΣ ΕΛΙΚΑΣ (SSB) ΚΑΙ ΔΙΠΛΩΝ ΘΡΑΥΣΜΑΤΩΝ (DSBs) ΠΟΥ ΜΠΟΡΟΥΝ ΝΑ ΕΠΕΛΘΟΥΝ ΕΠΕΙΤΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΙΟΝΤΙΖΟΥΣΑ ΑΚΤΙΝΟΒΟΛΙΑ (GEORGAKILAS, 2008).....	20
ΕΙΚΟΝΑ 13: ΑΠΕΙΚΟΝΙΣΗ ΤΗΣ ΔΙΑΔΙΚΑΣΙΑΣ ΚΑΤΑΣΚΕΥΗΣ ΚΑΙ ΑΝΑΓΝΩΣΗΣ ΠΕΙΡΑΜΑΤΟΣ ΜΙΚΡΟΣΥΣΤΟΙΧΙΩΝ ΓΙΑ ΤΗΝ ΜΕΛΕΤΗ ΤΗΣ ΓΟΝΙΔΙΑΚΗΣ ΕΚΦΡΑΣΗΣ (DUGGAN ET AL., 1999).....	24
ΕΙΚΟΝΑ 14: ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΕΝΟΣ ΘΗΚΟΓΡΑΜΜΑΤΟΣ.....	28
ΕΙΚΟΝΑ 15: Ο ΠΙΝΑΚΑΣ ΜΕ ΤΟ ΟΝΟΜΑ EXPRESSION (ΑΠΟ ΤΗΝ ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ MICROARRAY ΠΟΥ ΔΗΜΙΟΥΡΓΗΣΑΜΕ), ΕΧΕΙ 4 ΓΡΑΜΜΕΣ ΜΕ ΟΝΟΜΑΣΙΕΣ: GSM, FEATUREID, PLATFORM ΚΑΙ EXPRESSION. ΟΙ ΓΡΑΜΜΕΣ ΑΥΤΕΣ ΑΠΟΤΕΛΟΥΝ ΤΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΤΟΥ ΠΙΝΑΚΑ. ΔΕΞΙΑ ΑΠΟ ΤΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΕΙΝΑΙ Ο ΤΥΠΟΣ ΤΩΝ ΔΕΔΟΜΕΝΩΝ: VARCHAR, FLOAT ΚΑΙ ΣΕ ΠΑΡΕΝΘΕΣΗ ΕΙΝΑΙ ΤΟ ΠΕΔΙΟ ΟΡΙΣΜΟΥ ΤΟΥΣ. Ο ΠΙΝΑΚΑΣ ΕΧΕΙ ΒΑΘΜΟ 4.	30
ΕΙΚΟΝΑ 16: ΒΟΧΠΛΟΤ ΓΙΑ ΤΙΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΤΗΣ ΜΕΛΕΤΗΣ GSE8917 ΠΡΙΝ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ.	46
ΕΙΚΟΝΑ 17: ΒΟΧΠΛΟΤ ΜΕΤΑ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΓΙΑ ΤΗΝ ΣΕΙΡΑ GSE8917.....	46
ΕΙΚΟΝΑ 18: ΒΟΧΠΛΟΤ ΓΙΑ ΤΙΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΤΗΣ GSE12435 ΠΡΙΝ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ.....	47
ΕΙΚΟΝΑ 19: ΒΟΧΠΛΟΤ ΜΕΤΑ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΓΙΑ ΤΗΝ ΣΕΙΡΑ GSE12435.....	47
ΕΙΚΟΝΑ 20: ΒΟΧΠΛΟΤ ΓΙΑ ΤΙΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΤΗΣ GSE16935 ΠΡΙΝ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ.....	48
ΕΙΚΟΝΑ 21: ΒΟΧΠΛΟΤ ΜΕΤΑ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΓΙΑ ΤΗΝ ΣΕΙΡΑ GSE16935.....	48
ΕΙΚΟΝΑ 22: ΒΟΧΠΛΟΤ ΓΙΑ ΤΙΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΤΗΣ GSE23901 ΠΡΙΝ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ.....	49
ΕΙΚΟΝΑ 23: ΒΟΧΠΛΟΤ ΜΕΤΑ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΓΙΑ ΤΗΝ ΣΕΙΡΑ GSE23901.....	49
ΕΙΚΟΝΑ 24: Α) ΜΑΡΠΛΟΤ ΜΕ ΤΙΜΗ SPAN=0.4 ΚΑΙ Β) ΜΑΡΠΛΟΤ ΜΕ ΤΙΜΗ SPAN=0.75 ΓΙΑ ΤΟ ΙΔΙΟ SLIDE.....	51
ΕΙΚΟΝΑ 25: ΒΟΧΠΛΟΤ ΓΙΑ ΤΙΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΠΡΙΝ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΠΟΥ ΔΙΝΟΥΝ ΓΙΑ ΤΟ GSE6978.....	52
ΕΙΚΟΝΑ 26: ΒΟΧΠΛΟΤ ΜΕΤΑ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ ΓΙΑ ΤΗΝ ΣΕΙΡΑ GSE6978.....	52

ΕΙΚΟΝΑ 27: ΔΙΑΓΡΑΜΜΑ ΣΥΣΧΕΤΙΣΕΩΝ ΟΝΤΟΤΗΤΩΝ ΤΗΣ ΒΑΣΗΣ «MICROARRAY».....	56
ΕΙΚΟΝΑ 28: ΔΙΑΓΡΑΜΜΑΤΑ VENN ΜΕ CUTOFF=0.05 ΓΙΑ ΤΟ P-VALUE, ΓΙΑ ΔΕΙΓΜΑΤΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ ΜΕ ΚΡΙΤΗΡΙΟ ΤΟ ADJUSTED P-VALUE ΑΠΟ ΜΠΕΪΖΙΑΝΗ ΑΝΑΛΥΣΗ.....	70
ΕΙΚΟΝΑ 29: ΔΙΑΓΡΑΜΜΑΤΑ VENN ΜΕ CUTOFF=0.05 ΓΙΑ ΤΟ P-VALUE, ΓΙΑ ΔΕΙΓΜΑΤΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΧΑΜΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ ΜΕ ΚΡΙΤΗΡΙΟ ΤΟ ADJUSTED P-VALUE ΑΠΟ ΜΠΕΪΖΙΑΝΗ ΑΝΑΛΥΣΗ.....	70
ΕΙΚΟΝΑ 30: ΔΙΚΤΥΟ ΥΠΕΡΕΚΠΡΟΣΩΠΗΣΗΣ ΟΡΩΝ ΓΟΝΙΔΙΑΚΗΣ ΟΝΤΟΛΟΓΙΑΣ ΚΑΙ Ο ΣΥΣΧΕΤΙΣΜΟΣ ΑΥΤΩΝ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH.....	100
ΕΙΚΟΝΑ 31: ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΟΥ GAP JUNCTION PATHWAY ΠΟΥ ΠΡΟΕΚΥΨΕ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗ KEGG ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH.	101
ΕΙΚΟΝΑ 32: ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΟΥ P53 SIGNALING PATHWAY ΠΟΥ ΠΡΟΕΚΥΨΕ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗ KEGG ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH.....	102
ΕΙΚΟΝΑ 33 : ΔΙΚΤΥΟ ΥΠΕΡΕΚΠΡΟΣΩΠΗΣΗΣ ΟΡΩΝ ΓΟΝΙΔΙΑΚΗΣ ΟΝΤΟΛΟΓΙΑΣ ΚΑΙ Ο ΣΥΣΧΕΤΙΣΜΟΣ ΑΥΤΩΝ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΠΟΥ ΣΧΕΤΙΖΕΤΑΙ ΜΕ ΔΟΜΙΚΑ ΜΕΡΗ ΤΟΥ ΚΥΤΤΑΡΟΥ.....	105
ΕΙΚΟΝΑ 34: ΔΙΚΤΥΟ ΥΠΕΡΕΚΠΡΟΣΩΠΗΣΗΣ ΟΡΩΝ ΓΟΝΙΔΙΑΚΗΣ ΟΝΤΟΛΟΓΙΑΣ ΚΑΙ Ο ΣΥΣΧΕΤΙΣΜΟΣ ΑΥΤΩΝ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΧΩΡΙΣ ΔΙΟΡΘΩΣΗ ΠΟΛΛΑΠΛΩΝ ΕΡΩΤΗΜΑΤΩΝ. ΤΜΗΜΑ ΔΙΑΓΡΑΜΜΑΤΟΣ ΓΙΑ ΒΙΟΛΟΓΙΚΕΣ ΔΙΕΡΓΑΣΙΕΣ.	106
ΕΙΚΟΝΑ 35: ΔΙΚΤΥΟ ΥΠΕΡΕΚΠΡΟΣΩΠΗΣΗΣ ΟΡΩΝ ΓΟΝΙΔΙΑΚΗΣ ΟΝΤΟΛΟΓΙΑΣ ΚΑΙ Ο ΣΥΣΧΕΤΙΣΜΟΣ ΑΥΤΩΝ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΧΩΡΙΣ ΔΙΟΡΘΩΣΗ ΠΟΛΛΑΠΛΩΝ ΕΡΩΤΗΜΑΤΩΝ. ΤΜΗΜΑ ΔΙΑΓΡΑΜΜΑΤΟΣ ΓΙΑ ΜΟΡΙΑΚΕΣ ΛΕΙΤΟΥΡΓΙΕΣ.....	107
ΕΙΚΟΝΑ 36: ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΟΥ ADHERENS JUNCTION ΠΟΥ ΠΡΟΕΚΥΨΕ ΑΠΟ ΤΗΝ ΑΝΑΛΥΣΗ KEGG ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW.	108
ΕΙΚΟΝΑ 37: ΒΙΟΛΟΓΙΚΟ ΔΙΚΤΥΟ ΤΩΝ ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΣΜΕΝΩΝ ΓΟΝΙΔΙΩΝ ΟΠΩΣ ΠΡΟΕΚΥΨΑΝ ΑΠΟ ΤΟ ΣΥΝΟΛΟ ΤΩΝ ΔΕΙΓΜΑΤΩΝ ΠΟΥ ΕΚΤΕΘΗΚΑΝ ΣΕ ΧΑΜΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ.	111
ΕΙΚΟΝΑ 38: ΒΙΟΛΟΓΙΚΟ ΔΙΚΤΥΟ ΤΩΝ ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΣΜΕΝΩΝ ΓΟΝΙΔΙΩΝ ΟΠΩΣ ΠΡΟΕΚΥΨΑΝ ΑΠΟ ΤΟ ΣΥΝΟΛΟ ΤΩΝ ΔΕΙΓΜΑΤΩΝ ΠΟΥ ΕΚΤΕΘΗΚΑΝ ΣΕ ΥΨΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ.	112
ΕΙΚΟΝΑ 39: ΤΕΛΙΚΟ ΔΙΑΓΡΑΜΜΑ VENN ΟΠΟΥ ΣΥΝΔΥΑΖΟΝΤΑΙ ΤΑ ΓΟΝΙΔΙΑ ΠΟΥ ΞΕΧΩΡΙΣΑΝ ΣΕ ΑΥΤΗ ΤΗΝ ΕΡΓΑΣΙΑ ΜΕ ΤΑ ΓΟΝΙΔΙΑ ΑΠΟ ΤΗΝ ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΑΙ ΤΗΝ ΛΙΣΤΑ ΓΟΝΙΔΙΩΝ ΠΟΥ ΣΧΕΤΙΖΟΝΤΑΙ ΜΕ ΤΗΝ ΔΗΜΙΟΥΡΓΙΑ ΚΑΡΚΙΝΟΥ. ΤΟ ΟΡΙΟ ΓΙΑ ΤΙΣ ΧΑΜΗΛΕΣ ΔΟΣΕΙΣ ΕΙΝΑΙ ΤΑ 0.5Gy.....	122
ΠΙΝΑΚΑΣ 1: ΥΠΟΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ ΓΙΑ ΤΗΝ ΕΦΑΡΜΟΓΗ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗΣ ΠΟΣΟΣΤΗΜΟΡΙΟΥ.	27
ΠΙΝΑΚΑΣ 2: ΤΕΛΙΚΗ ΤΑΞΙΝΟΜΗΣΗ ΓΙΑ ΚΑΘΕ ΔΕΙΓΜΑ ΜΕ ΕΦΑΡΜΟΓΗ ΤΗΣ ΜΕΘΟΔΟΥ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗΣ ΠΟΣΟΣΤΗΜΟΡΙΟΥ.	27
ΠΙΝΑΚΑΣ 3: ΑΝΑΛΥΤΙΚΑ ΣΤΟΙΧΕΙΑ ΓΙΑ ΤΙΣ ΜΕΛΕΤΕΣ ΠΟΥ ΕΠΙΛΕΧΘΗΚΑΝ ΚΑΙ ΑΝΑΛΥΘΗΚΑΝ ΜΕ ΤΙΣ ΑΝΤΙΣΤΟΙΧΕΣ ΠΛΑΤΦΟΡΜΕΣ ΠΟΥ ΤΟΥΣ ΑΝΤΙΣΤΟΙΧΟΥΝ.	40
ΠΙΝΑΚΑΣ 4: ΛΙΣΤΑ ΔΕΙΓΜΑΤΩΝ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΣΤΗΝ ΠΑΡΟΥΣΑ ΕΡΓΑΣΙΑ ΜΕ ΤΗΝ ΑΝΤΙΣΤΟΙΧΗ ΚΑΤΑΣΤΑΣΗ (CONTROL-LOW-HIGH) ΚΑΙ ΤΗΝ ΑΝΤΙΣΤΟΙΧΗ ΜΕΛΕΤΗ ΣΤΗΝ ΟΠΟΙΑ ΑΝΗΚΟΥΝ.....	43
ΠΙΝΑΚΑΣ 5: ΜΟΡΦΗ ΑΡΧΙΚΟΥ ΑΡΧΕΙΟΥ ΑΝΑΓΝΩΣΗΣ ΟΠΩΣ ΕΙΣΗΧΘΕΙ ΣΤΟ RSTUDIO ΓΙΑ ΤΙΣ ΣΤΑΘΜΙΣΜΕΝΕΣ ΕΚΦΡΑΣΕΙΣ ΤΗΣ ΣΕΙΡΑΣ GSE8917 ΓΙΑ ΤΟΝ ΥΠΟΛΟΓΙΣΜΟ ΤΩΝ P-VALUES.....	63
ΠΙΝΑΚΑΣ 6: ΕΝΔΕΙΚΤΙΚΗ ΣΥΓΚΡΙΣΗ ΕΚΦΡΑΣΗΣ ΟΡΙΣΜΕΝΩΝ ΓΟΝΙΔΙΩΝ ΜΙΑΣ ΣΕΙΡΑΣ ΠΕΙΡΑΜΑΤΩΝ ΠΟΥ ΑΝΗΚΟΥΝ ΣΤΗΝ ΙΔΙΑ ΜΕΛΕΤΗ ΚΑΙ ΣΤΟ ΙΔΙΟ GROUP ΟΡΩΝ ΓΙΑ ΔΥΟ ΚΑΤΑΣΤΑΣΕΙΣ ΧΡΗΣΙΜΟΠΟΙΩΝΤΑΣ ΤΙΣ ΑΝΤΙΣΤΟΙΧΕΣ ΤΙΜΕΣ ΤΩΝ ADJUSTED P-VALUES.	64
ΠΙΝΑΚΑΣ 7: ΑΡΙΘΜΟΣ ΓΟΝΙΔΙΩΝ ΠΟΥ ΑΝΑΔΕΙΧΘΗΚΑΝ ΩΣ ΣΤΑΤΙΣΤΙΚΩΣ ΣΗΜΑΝΤΙΚΑ ΑΠΟ ΤΗΝ ΜΕΤΑ-ΑΝΑΛΥΣΗ ΒΑΣΗ ΤΗΣ ΤΙΜΗΣ P-VALUE<0.01 ΓΙΑ ΚΑΘΕ ΣΥΝΟΛΟ ΓΟΝΙΔΙΩΝ.....	71
ΠΙΝΑΚΑΣ 8: ΓΟΝΙΔΙΑ ΠΟΥ ΞΕΧΩΡΙΣΑΝ ΩΣ ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΣΜΕΝΑ ΜΕΣΩ ΜΕΤΑ-ΑΝΑΛΥΣΗΣ, ΜΕ ΒΑΣΗ ΤΙΜΗ P-VALUE<0.01, ΕΠΕΙΤΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ(>0.5Gy) ΣΤΟ ΔΙΑΣΤΗΜΑ 4-16 ΩΡΩΝ. ΣΥΜΠΕΡΙΛΑΜΒΑΝΕΤΑΙ ΤΟ ΑΝΤΙΣΤΟΙΧΟ HGNC ΣΥΜΒΟΛΟ ΜΕ ΤΗΝ ΑΝΤΙΣΤΟΙΧΗ ΠΕΡΙΓΡΑΦΗ ΚΑΙ ΟΙ ΑΚΡΙΒΕΙΣ ΤΙΜΕΣ ΤΟΥ ΥΠΟΛΟΓΙΣΜΕΝΟΥ P-VALUE ΚΑΙ ΤΗΣ ΔΙΑΦΟΡΑΣ ΣΤΗΝ ΕΚΦΡΑΣΗ.	71

ΠΙΝΑΚΑΣ 9: ΓΟΝΙΔΙΑ ΠΟΥ ΞΕΧΩΡΙΣΑΝ ΩΣ ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΣΜΕΝΑ ΜΕΣΩ ΜΕΤΑ-ΑΝΑΛΥΣΗΣ, ΜΕ ΒΑΣΗ ΤΙΜΗ P-VALUE<0.01, ΕΠΕΙΤΑ ΑΠΟ ΕΚΘΕΣΗ ΣΕ ΧΑΜΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ(<0.5Gy) ΣΤΟ ΔΙΑΣΤΗΜΑ 4-16 ΩΡΩΝ. ΣΥΜΠΕΡΙΛΑΜΒΑΝΕΤΑΙ ΤΟ ΑΝΤΙΣΤΟΙΧΟ HGNC ΣΥΜΒΟΛΟ ΜΕ ΤΗΝ ΑΝΤΙΣΤΟΙΧΗ ΠΕΡΙΓΡΑΦΗ ΚΑΙ ΟΙ ΑΚΡΙΒΕΙΣ ΤΙΜΕΣ ΤΟΥ ΥΠΟΛΟΓΙΣΜΕΝΟΥ P-VALUE ΚΑΙ ΤΗΣ ΔΙΑΦΟΡΑΣ ΣΤΗΝ ΕΚΦΡΑΣΗ.....	89
ΠΙΝΑΚΑΣ 10: ΑΡΙΘΜΟΣ ΓΟΝΙΔΙΩΝ ΠΟΥ ΑΝΤΙΣΤΟΙΧΗΘΗΚΑΝ ΜΟΝΑΔΙΚΑ ΑΠΟ ΕΚΕΙΝΑ ΠΟΥ ΕΞΗΓΑΜΕ ΑΠΟ ΤΗΝ ΜΕΤΑ-ΑΝΑΛΥΣΗ ΚΑΙ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΑΠΟ ΤΟ ΣΥΣΤΗΜΑ ΓΙΑ ΤΙΣ ΑΝΑΛΥΣΕΙΣ ΓΙΑ ΚΑΘΕ ΣΥΝΟΛΟ ΓΟΝΙΔΙΩΝ.	98
ΠΙΝΑΚΑΣ 11: ΑΝΑΛΥΣΕΙΣ ΠΟΥ ΕΔΩΞΑΝ ΑΠΟΤΕΛΕΣΜΑΤΑ ΓΙΑ ΚΑΘΕ ΣΥΝΟΛΟ ΓΟΝΙΔΙΩΝ. ΣΤΙΣ ΠΑΡΕΝΘΕΣΕΙΣ ΑΝΑΓΡΑΦΕΤΑΙ ΤΟ ΕΠΙΠΕΔΟ ΣΗΜΑΝΤΙΚΟΤΗΤΑΣ ΚΑΙ Ο ΔΕΣΤΕΡΙΣΚΟΣ ΣΗΜΑΤΟΔΟΤΕΙ ΤΗΝ ΑΠΟΥΣΙΑ ΔΙΟΡΘΩΣΗΣ ΠΟΛΛΑΠΛΩΝ ΕΡΩΤΗΜΑΤΩΝ. ΤΕΛΟΣ ΟΙ ΑΝΑΛΥΣΕΙΣ TFT ΚΑΙ MICRORNA ΕΓΙΝΑΝ ΑΦΟΥ ΔΙΑΧΩΡΙΣΑΜΕ ΤΑ UP ΚΑΙ DOWN REGULATED ΓΟΝΙΔΙΑ.	98
ΠΙΝΑΚΑΣ 12:ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH (ΕΙΚΟΝΑ 30).....	99
ΠΙΝΑΚΑΣ 13: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ KEGG ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH (ΕΙΚΟΝΑ 31, ΕΙΚΟΝΑ 32).	99
ΠΙΝΑΚΑΣ 14: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ PATHWAY COMMONS ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH.....	99
ΠΙΝΑΚΑΣ 15: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ WIKI PATHWAYS ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH.....	99
ΠΙΝΑΚΑΣ 16: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ TRANSCRIPTION FACTOR TARGET ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH ΓΙΑ ΤΑ DOWN-REGULATED ΓΟΝΙΔΙΑ. ...	103
ΠΙΝΑΚΑΣ 17: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ MICRORNA TARGET ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-HIGH ΓΙΑ ΤΑ DOWN-REGULATED ΓΟΝΙΔΙΑ	103
ΠΙΝΑΚΑΣ 18: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΣΧΕΤΙΚΑ ΜΕ ΤΑ ΜΕΡΗ ΤΟΥ ΚΥΤΤΑΡΟΥ (ΕΙΚΟΝΑ 33).	104
ΠΙΝΑΚΑΣ 19: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΧΩΡΙΣ ΔΙΟΡΘΩΣΗ ΠΟΛΛΑΠΛΩΝ ΕΡΩΤΗΜΑΤΩΝ ΣΤΟ ΚΟΜΜΑΤΙ ΤΩΝ ΒΙΟΛΟΓΙΚΩΝ ΔΙΕΡΓΑΣΙΩΝ (ΕΙΚΟΝΑ 34).	104
ΠΙΝΑΚΑΣ 20: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ GO ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΧΩΡΙΣ ΔΙΟΡΘΩΣΗ ΠΟΛΛΑΠΛΩΝ ΕΡΩΤΗΜΑΤΩΝ ΣΤΟ ΚΟΜΜΑΤΙ ΤΩΝ ΜΟΡΙΑΚΩΝ ΛΕΙΤΟΥΡΓΙΩΝ (ΕΙΚΟΝΑ 35).	107
ΠΙΝΑΚΑΣ 21: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ WIKI PATHWAYS ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW.	108
ΠΙΝΑΚΑΣ 22: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ PATHWAY COMMONS ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW.	108
ΠΙΝΑΚΑΣ 23: ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ KEGG ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW (ΕΙΚΟΝΑ 36).....	108
ΠΙΝΑΚΑΣ 24: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ TRANSCRIPTION FACTOR TARGET ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΓΙΑ ΤΑ DOWN-REGULATED ΓΟΝΙΔΙΑ.	109
ΠΙΝΑΚΑΣ 25: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ TRANSCRIPTION FACTOR TARGET ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΓΙΑ ΤΑ UP-REGULATED ΓΟΝΙΔΙΑ.	109
ΠΙΝΑΚΑΣ 26: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ MICRORNA TARGET ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΓΙΑ ΤΑ DOWN-REGULATED ΓΟΝΙΔΙΑ.....	109
ΠΙΝΑΚΑΣ 27: ΚΥΡΙΟΤΕΡΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ MICRORNA TARGET ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ CONTROL-LOW ΓΙΑ ΤΑ UP-REGULATED ΓΟΝΙΔΙΑ.	109
ΠΙΝΑΚΑΣ 28: ΑΡΙΘΜΟΣ ΓΟΝΙΔΙΩΝ ΠΟΥ ΑΝΤΙΣΤΟΙΧΗΘΗΚΑΝ ΣΤΗΝ ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ.	110
ΠΙΝΑΚΑΣ 29: ΣΤΑΤΙΣΤΙΚΑ ΣΤΟΙΧΕΙΑ ΓΙΑ ΤΗΝ ΑΠΕΙΚΟΝΙΣΗ ΤΩΝ ΒΙΟΛΟΓΙΚΩΝ ΔΙΚΤΥΩΝ ΑΠΟ ΤΟ STRING.	111
ΠΙΝΑΚΑΣ 30: PATHWAY ENRICHMENT ΓΙΑ ΤΟ ΣΥΝΟΛΟ ΓΟΝΙΔΙΩΝ ΠΟΥ ΠΡΟΕΚΥΨΑΝ ΑΠΟ ΔΕΙΓΜΑΤΑ ΤΑ ΟΠΟΙΑ ΕΚΤΕΘΗΚΑΝ ΣΕ ΧΑΜΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ (<0.5Gy).	111
ΠΙΝΑΚΑΣ 31: ΣΥΝΟΨΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΓΙΑ ΤΗΝ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ (>0.5Gy).	113
ΠΙΝΑΚΑΣ 32: ΣΥΝΟΨΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΓΙΑ ΤΗΝ ΕΚΘΕΣΗ ΣΕ ΧΑΜΗΛΗ ΔΟΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ (<0.5Gy).	114
ΠΙΝΑΚΑΣ 33: ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΖΟΜΕΝΑ ΓΟΝΙΔΙΑ ΠΟΥ ΠΡΟΚΥΠΤΟΥΝ ΑΠΟ ΤΗΝ ΒΙΒΛΙΟΓΡΑΦΙΑ ΓΙΑ ΤΙΣ ΥΨΗΛΕΣ ΔΟΣΕΙΣ (>0.5Gy).	118
ΠΙΝΑΚΑΣ 34: ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΖΟΜΕΝΑ ΓΟΝΙΔΙΑ ΠΟΥ ΠΡΟΚΥΠΤΟΥΝ ΑΠΟ ΤΗΝ ΒΙΒΛΙΟΓΡΑΦΙΑ ΓΙΑ ΤΙΣ ΧΑΜΗΛΕΣ ΔΟΣΕΙΣ (<0.5Gy).....	119
ΠΙΝΑΚΑΣ 35: ΓΟΝΙΔΙΑ ΠΟΥ ΒΡΕΘΗΚΑΝ ΚΟΙΝΑ ΜΕΤΑΞΥ ΤΗΣ ΜΕΛΕΤΗΣ ΜΑΣ ΓΙΑ ΤΗΝ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΕΣ ΔΟΣΕΙΣ ΜΕ ΤΗΝ ΒΙΒΛΙΟΓΡΑΦΙΑ.	123

ΠΙΝΑΚΑΣ 36: ΓΟΝΙΔΙΑ ΠΟΥ ΒΡΕΘΗΚΑΝ ΣΤΗΝ ΜΕΛΕΤΗ ΜΑΣ ΚΟΙΝΑ ΓΙΑ ΤΗΝ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΕΣ ΚΑΙ ΧΑΜΗΛΕΣ ΔΟΣΕΙΣ.....	123
ΠΙΝΑΚΑΣ 37: ΓΟΝΙΔΙΑ ΠΟΥ ΕΝΤΟΠΙΣΑΜΕ ΑΠΟ ΤΗΝ ΒΙΒΛΙΟΓΡΑΦΙΑ ΚΑΙ ΣΥΜΠΙΠΟΥΝ ΜΕ ΤΗΝ ΛΙΣΤΑ ΓΙΑ ΤΑ ΓΟΝΙΔΙΑ ΠΟΥ ΣΧΕΤΙΖΟΝΤΑΙ ΜΕ ΤΗΝ ΕΜΦΑΝΙΣΗ ΚΑΡΚΙΝΟΥ	124
ΠΙΝΑΚΑΣ 38: ΓΟΝΙΔΙΑ ΠΟΥ ΞΕΧΩΡΙΣΑΜΕ ΓΙΑ ΤΗΝ ΕΚΘΕΣΗ ΣΕ ΥΨΗΛΕΣ ΔΟΣΕΙΣ ΚΑΙ ΣΥΜΠΙΠΟΥΝ ΜΕ ΤΗΝ ΛΙΣΤΑ ΓΙΑ ΤΑ ΓΟΝΙΔΙΑ ΠΟΥ ΣΧΕΤΙΖΟΝΤΑΙ ΜΕ ΤΗΝ ΕΜΦΑΝΙΣΗ ΚΑΡΚΙΝΟΥ.....	124
ΠΙΝΑΚΑΣ 39: ΓΟΝΙΔΙΑ ΠΟΥ ΞΕΧΩΡΙΣΑΜΕ ΓΙΑ ΤΗΝ ΕΚΘΕΣΗ ΣΕ ΧΑΜΗΛΕΣ ΔΟΣΕΙΣ ΚΑΙ ΣΥΜΠΙΠΟΥΝ ΜΕ ΤΗΝ ΛΙΣΤΑ ΓΙΑ ΤΑ ΓΟΝΙΔΙΑ ΠΟΥ ΣΧΕΤΙΖΟΝΤΑΙ ΜΕ ΤΗΝ ΕΜΦΑΝΙΣΗ ΚΑΡΚΙΝΟΥ.....	124
ΕΙΣΩΣΗ 1: ΓΕΝΙΚΗ ΜΟΡΦΗ ΤΗΣ ΕΙΣΩΣΗΣ ΒΕΤΗΕ-BLOCH, ΟΠΟΥ $(1/4\pi e_0)$ ΣΤΑΘΕΡΗ ΠΟΣΟΤΗΤΑ ΑΠΟ ΤΟΝ ΝΟΜΟ COULOMB, Μ, Ζ ΚΑΙ U ΚΑΘΟΡΙΖΟΝΤΑΙ ΑΠΟ ΤΗΝ ΑΚΤΙΝΟΒΟΛΙΑ (ΜΑΖΑ, ΑΤΟΜΙΚΟΣ ΑΡΙΘΜΟΣ, Π.Χ. $Z=1$ ΓΙΑ ΠΡΩΤΟΝΙΑ, ΤΑΧΥΤΗΤΑ), Ε ΤΟ ΦΟΡΤΙΟ ΤΟΥ ΗΛΕΚΤΡΟΝΙΟΥ ΚΑΙ Ι ΤΟ ΔΥΝΑΜΙΚΟ ΙΟΝΤΙΣΜΟΥ.	7
ΕΙΣΩΣΗ 2: ΕΙΣΩΣΗ ΣΧΕΤΙΚΙΣΤΙΚΗΣ ΟΡΜΗΣ ΜΕ Β ΚΑΙ Γ ΣΧΕΤΙΚΙΣΤΙΚΕΣ ΣΤΑΘΕΡΕΣ, Μ Η ΜΑΖΑ ΤΟΥ ΣΩΜΑΤΙΔΙΟΥ ΚΑΙ C Η ΤΑΧΥΤΗΤΑ ΤΟΥ ΦΩΤΟΣ.....	7
ΕΙΣΩΣΗ 3: ΓΡΑΜΜΙΚΗ ΜΕΤΑΒΙΒΑΣΗ ΕΝΕΡΓΕΙΑΣ (LINEAR ENERGY TRANSFER).	9
ΕΙΣΩΣΗ 4: ΥΠΟΛΟΓΙΖΟΜΕΝΗ ΕΚΦΡΑΣΗ ΓΙΑ ΕΝΑ ΧΡΩΜΑ, ΣΕ ΕΝΑ SPOT ΜΕΣΟΥ ΤΗΣ ΜΕΘΟΔΟΥ NORMEXR. ΟΠΟΥ Rf Η ΕΝΤΑΣΗ ΕΚΦΡΑΣΗΣ ΣΤΟ ΠΡΟΣΚΗΝΙΟ (FOREGROUND), Rb Η ΕΝΤΑΣΗ ΣΤΟ ΥΠΟΒΑΘΡΟ (BACKGROUND), S Η ΦΩΤΕΙΝΟΤΗΤΑ ΠΟΥ ΑΝΤΙΣΤΟΙΧΕΙ ΣΤΗΝ ΕΝΤΑΣΗ ΤΗΣ ΠΡΑΓΜΑΤΙΚΗΣ ΕΚΦΡΑΣΗΣ ΚΑΙ ΑΚΟΛΟΥΘΕΙ ΕΚΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ ΕΝΩ B ΕΙΝΑΙ Η ΕΝΤΑΣΗ ΥΠΟΒΑΘΡΟΥ ΠΟΥ ΔΕΝ ΣΥΜΠΕΡΙΛΑΜΒΑΝΕΤΑΙ ΣΤΟ Rb, Η ΟΠΟΙΑ ΑΚΟΛΟΥΘΕΙ ΕΚΘΕΤΙΚΗ ΚΑΤΑΝΟΜΗ.	26
ΕΙΣΩΣΗ 5: ΤΕΛΙΚΕΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΓΙΑ ΤΑ ΔΥΟ ΧΡΩΜΑΤΑ ΕΠΕΙΤΑ ΑΠΟ ΔΙΟΡΘΩΣΗ ΥΠΟΒΑΘΡΟΥ.	26
ΕΙΣΩΣΗ 6: ΤΕΛΙΚΕΣ ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΕΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΓΙΑ ΔΙΧΡΩΜΑΤΙΚΗ ΜΙΚΡΟΣΥΣΤΟΙΧΙΑ.	26
ΕΙΣΩΣΗ 7: ΜΕΣΟΣ ΟΡΟΣ ΕΚΦΡΑΣΗΣ ΑΠΟ ΟΛΑ ΤΑ ΔΕΙΓΜΑΤΑ ΕΠΕΙΤΑ ΑΠΟ ΤΑΞΙΝΟΜΗΣΗ.....	27
ΕΙΣΩΣΗ 8: ΜΗ-ΤΥΠΟΠΟΙΗΜΕΝΗ ΔΙΑΦΟΡΑ ΜΕΣΩΝ ΤΙΜΩΝ (UNSTANDARDIZED MEAN DIFFERENCE).	35
ΕΙΣΩΣΗ 9: ΤΥΠΟΠΟΙΗΜΕΝΗ ΔΙΑΦΟΡΑ ΜΕΣΩΝ ΤΙΜΩΝ (STANDARDIZED MEAN DIFFERENCE).	35
ΕΙΣΩΣΗ 10: ΕΝΟΠΟΙΗΜΕΝΗ ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ ΓΙΑ ΔΥΟ ΚΑΤΑΣΤΑΣΕΙΣ. ΤΑ Ν ΥΠΟΔΗΛΩΝΟΥΝ ΤΟΝ ΑΡΙΘΜΟ ΤΩΝ ΔΕΙΓΜΑΤΩΝ ΓΙΑ ΤΙΣ ΔΥΟ ΚΑΤΑΣΤΑΣΕΙΣ.	35
ΕΙΣΩΣΗ 11: ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ ΓΙΑ ΚΑΘΕ ΚΑΤΑΣΤΑΣΗ ΜΕΜΟΝΩΜΕΝΑ, ΟΠΟΥ Ο ΠΑΡΟΝΟΜΑΣΤΗΣ ΕΚΦΡΑΖΕΙ ΤΟΝ ΒΑΘΜΟ ΕΛΕΥΘΕΡΙΑΣ ΤΗΣ ΜΕΛΕΤΗΣ ΚΑΘΩΣ ΤΟ Ν ΥΠΟΔΗΛΩΝΕΙ ΤΟΝ ΑΡΙΘΜΟ ΤΩΝ ΔΕΙΓΜΑΤΩΝ.	35
ΕΙΣΩΣΗ 12: ΑΠΟΚΛΙΣΗ ΤΗΣ ΤΥΠΟΠΟΙΗΜΕΝΗΣ ΜΕΣΗΣ ΤΙΜΗΣ ΤΗΣ ΔΙΑΦΟΡΑΣ (STANDARDIZED MEAN DIFFERENCE DEVIATION) ΓΙΑ ΜΙΑ ΜΕΛΕΤΗ.	35
ΕΙΣΩΣΗ 13: ΔΙΟΡΘΩΤΙΚΟΣ ΠΑΡΑΓΟΝΤΑΣ J ΓΙΑ ΜΕΤΑΤΡΟΠΗ ΣΤΟ G ΤΟΥ HEDGES, ΟΠΟΥ DF Ο ΒΑΘΜΟΣ ΕΛΕΥΘΕΡΙΑΣ $(n1+n2-2)$ ΓΙΑ ΕΝΑ ΓΟΝΙΔΙΟ.	35
ΕΙΣΩΣΗ 14: ΜΕΓΕΘΟΣ ΕΠΙΔΡΑΣΗΣ ΚΑΘΕ ΜΕΛΕΤΗΣ (HEDGE'S G).	35
ΕΙΣΩΣΗ 15: ΑΠΟΚΛΙΣΗ ΤΟΥ ΜΕΓΕΘΟΥΣ ΕΠΙΔΡΑΣΗΣ.	36
ΕΙΣΩΣΗ 16: ΤΟ WEIGHT ΚΑΘΕ ΜΕΛΕΤΗΣ ΕΠΕΙΤΑ ΑΠΟ ΜΕΤΑΤΡΟΠΗ ΣΤΟ G ΤΟΥ HEDGES ΚΑΙ ΧΡΗΣΗ ΑΥΤΟΥ ΩΣ ΜΕΓΕΘΟΣ ΕΠΙΔΡΑΣΗΣ.	36
ΕΙΣΩΣΗ 17: ΥΠΟΛΟΓΙΖΕΤΑΙ ΩΣ T-SQUARED Η ΔΙΑΚΥΜΑΝΣΗ ΤΩΝ ΜΕΓΕΘΩΝ ΕΠΙΔΡΑΣΗΣ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΣΤΟΝ ΠΛΗΘΥΣΜΟ ΤΩΝ ΜΕΛΕΤΩΝ.	36
ΕΙΣΩΣΗ 18: ΝΕΑ ΑΠΟΚΛΙΣΗ ΤΟΥ ΜΕΓΕΘΟΥΣ ΕΠΙΔΡΑΣΗΣ ΣΕ ΣΥΝΔΥΑΣΜΟ ΜΕ ΤΟ T-SQUARED.	36
ΕΙΣΩΣΗ 19: ΤΟ ΝΕΟ WEIGHT ΚΑΘΕ ΜΕΛΕΤΗΣ ΜΕ ΤΗΝ ΕΠΙΔΡΑΣΗ ΤΟΥ T-SQUARED.	36
ΕΙΣΩΣΗ 20: ΤΕΛΙΚΗ ΤΙΜΗ ΔΙΑΦΟΡΑΣ ΕΚΦΡΑΣΗΣ ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ ΔΥΟ ΚΑΤΑΣΤΑΣΕΩΝ ΓΙΑ ΕΝΑ ΓΟΝΙΔΙΟ ΑΠΟ ΟΛΕΣ ΤΙΣ ΜΕΛΕΤΕΣ ΣΤΙΣ ΟΠΟΙΕΣ ΕΜΦΑΝΙΣΤΗΚΕ.	37
ΕΙΣΩΣΗ 21: ΑΠΟΚΛΙΣΗ ΤΗΣ ΤΕΛΙΚΗΣ ΤΙΜΗΣ ΔΙΑΦΟΡΑΣ ΕΚΦΡΑΣΗΣ ΓΙΑ ΕΝΑ ΓΟΝΙΔΙΟ ΣΕ ΜΙΑ ΣΥΓΚΡΙΣΗ.....	37

ΕΙΣΩΣΗ 22: ΤΥΠΙΚΟ ΣΦΑΛΜΑ (STANDARD ERROR) ΓΙΑ ΤΗΝ ΤΙΜΗ ΤΗΣ ΔΙΑΦΟΡΑΣ ΕΚΦΡΑΣΗΣ, ΓΙΑ ΤΗΝ ΣΥΓΚΡΙΣΗ ΕΝΟΣ ΓΟΝΙΔΙΟΥ, ΓΙΑ ΟΛΕΣ ΤΙΣ ΜΕΛΕΤΕΣ ΟΠΟΥ ΕΜΦΑΝΙΣΤΗΚΕ.....	37
ΕΙΣΩΣΗ 23: ΥΠΟΛΟΓΙΣΜΟΣ ΤΗΣ ΤΙΜΗΣ Z ΓΙΑ ΤΗΝ ΜΗΔΕΝΙΚΗ ΥΠΟΘΕΣΗ.....	37
ΕΙΣΩΣΗ 24: ΥΠΟΛΟΓΙΣΜΟΣ P-VALUE, ΟΠΟΥ $\Phi(Z)$ Η ΤΥΠΙΚΗ ΚΑΝΟΝΙΚΗ ΑΘΡΟΙΣΤΙΚΗ ΚΑΤΑΝΟΜΗ, Η ΟΠΟΙΑ ΕΦΑΡΜΟΖΕΤΑΙ ΣΤΟ EXCEL ΩΣ NORMSDIST(Z) ΘΕΩΡΩΝΤΑΣ ΜΕΣΟ (MEAN) ΤΟ ΜΗΔΕΝ ΚΑΙ ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ 1.....	37
ΕΙΣΩΣΗ 25: ΚΑΝΟΝΙΚΟΠΟΙΗΜΕΝΕΣ ΤΙΜΕΣ ΕΚΦΡΑΣΗΣ ΓΙΑ ΔΙΧΡΩΜΑΤΙΚΗ ΜΙΚΡΟΣΥΣΤΟΙΧΙΑ.....	51
SCRIPT 1: ENSG.EDIT.PHP ΓΙΑ ΤΗΝ ΔΙΑΤΗΡΗΣΗ ΤΩΝ ΑΠΑΡΑΙΤΗΤΩΝ ΠΛΗΡΟΦΟΡΙΩΝ ΓΙΑ ΤΟΝ ΠΙΝΑΚΑ ENSG_HGNC.....	57
SCRIPT 2: ENST_SEQ.PHP ΓΙΑ ΤΗΝ ΔΙΑΧΕΙΡΙΣΗ ΤΩΝ ΑΛΛΗΛΟΥΧΙΩΝ.....	57
SCRIPT 3: SEQ2ID.PHP. ΚΩΔΙΚΟΠΟΙΗΣΗ ΑΛΛΗΛΟΥΧΙΑΣ ΜΕ ΣΚΟΠΟ ΤΗΝ ΓΡΗΓΟΡΟΤΕΡΗ ΛΕΙΤΟΥΡΓΙΑ ΤΩΝ ΕΝΤΟΛΩΝ ΤΗΣ SQL.....	58
SCRIPT 4: PROBES.PHP. ΕΠΕΞΕΡΓΑΣΙΑ ΟΛΩΝ ΤΩΝ MANIFEST FILES ΤΩΝ ΠΛΑΤΦΟΡΜΩΝ ΚΑΙ ΕΞΑΓΩΓΗ ΤΩΝ ΑΠΑΡΑΙΤΗΤΩΝ ΣΤΟΙΧΕΙΩΝ (PROBEID ΚΑΙ ΑΝΤΙΣΤΟΙΧΗ ΑΚΟΛΟΥΘΙΑ).....	59
SCRIPT 5: P_VALUE.CUTOFF ΓΙΑ ΕΠΙΛΟΓΗ ΤΩΝ ΣΤΑΤΙΣΤΙΚΑ ΣΗΜΑΝΤΙΚΩΝ ΔΙΑΦΟΡΙΚΑ ΕΚΦΡΑΣΜΕΝΩΝ ΓΟΝΙΔΙΩΝ.....	67
SCRIPT 6: IONISING_HOURS.PHP ΓΙΑ ΤΟΝ ΔΙΑΧΩΡΙΣΜΟ ΤΩΝ ΔΕΙΓΜΑΤΩΝ ΜΕ ΒΑΣΗ ΤΗΝ ΩΡΑ ΠΟΥ ΣΥΛΛΕΧΘΗΚΑΝ.....	68
ΓΡΑΦΗΜΑ 1: ΑΝΑΣΧΕΤΙΚΗ ΙΣΧΥΣ ΜΙΟΝΙΩΝ ΣΕ ΧΑΛΚΟ ΣΕ ΣΧΕΣΗ ΜΕ ΤΗΝ ΣΧΕΤΙΚΙΣΤΙΚΗ ΟΡΜΗ.....	8
ΓΡΑΦΗΜΑ 2: ΚΑΜΠΥΛΗ ΕΝΑΠΟΘΕΣΗΣ ΕΝΕΡΓΕΙΑΣ ΩΣ ΠΡΟΣ ΒΑΘΟΣ ΕΙΣΧΩΡΗΣΗΣ ΔΙΑΦΟΡΩΝ ΑΚΤΙΝΟΒΟΛΙΩΝ. ΞΕΧΩΡΙΖΕΙ Η ΠΕΡΙΠΤΩΣΗ ΠΡΩΤΟΝΙΩΝ ΚΑΙ ΙΟΝΤΩΝ ΑΝΘΡΑΚΑ ΟΠΟΥ ΣΧΗΜΑΤΙΖΕΤΑΙ Η ΧΑΡΑΚΤΗΡΙΣΤΙΚΗ ΚΑΜΠΥΛΗ BRAGG.....	8
ΓΡΑΦΗΜΑ 3: ΚΑΜΠΥΛΗ ΕΚΠΟΜΠΗΣ ΩΣ ΠΡΟΣ ΤΟ ΠΑΧΟΣ ΤΟΥ ΑΠΟΡΡΟΦΗΤΗ. ΠΑΡΑΤΗΡΕΙΤΑΙ ΤΟ ΦΑΙΝΟΜΕΝΟ ΤΟΥ ΣΤΡΑΓΓΑΛΙΣΜΟΥ, ΟΠΟΥ MEAN RANGE: 50% ΤΩΝ ΣΩΜΑΤΙΔΙΩΝ ΠΟΥ ΕΚΠΕΜΦΘΗΚΑΝ ΚΑΙ EXTRAPOLATED RANGE: ΤΟ ΠΑΧΟΣ ΟΠΟΥ ΑΝΑΜΕΝΕΤΑΙ ΜΗΔΕΝΙΚΗ ΕΚΠΟΜΠΗ-ΑΠΟΡΡΟΦΗΣΗ. 9	9

Περιεχόμενα

Περιεχόμενα	1
1 ΙΟΝΤΙΖΟΥΣΕΣ ΑΚΤΙΝΟΒΟΛΙΕΣ	3
1.1 Εκπομπή & Ιδιότητες Ιοντιζουσών Ακτινοβολιών	3
1.2 Χαρακτηριστικά της Ακτινοβολίας	7
1.3 Το Χημικό Στάδιο Αλληλεπίδρασης	10
Το Βιολογικό Στάδιο Αλληλεπίδρασης	12
1.3.1 Βιολογικές Αποκρίσεις	12
1.3.2 Δομή DNA και Κυτταρικός Κύκλος	14
1.3.3 Βλάβες DNA	15
1.3.4 Χαμηλές Δόσεις Ακτινοβολίας	15
1.3.5 Υψηλές Δόσεις Ακτινοβολίας	19
1.3.6 Ο ρόλος του LET–Ομαδοποιημένες Βλάβες	19
2 ΒΙΟΠΛΗΡΟΦΟΡΙΚΕΣ ΜΕΘΟΔΟΙ ΕΠΕΞΕΡΓΑΣΙΑΣ	22
2.1 Μικροσυτοιχίες	22
2.2 Πακέτο limma	25
2.2.1 Η Μέθοδος “normexp” και η Τιμή offset	25
2.2.2 Κανονικοποίηση Ποσοστημορίου	26
2.2.3 Έλεγχος με Χρήση Θηκογράμματος	28
3 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ	29
3.1 Μοντέλο Διασύνδεσης Δεδομένων	29
3.2 Συσχετίσεις	31
3.3 Διάγραμμα Συσχετίσεων Οντοτήτων (ERD)	31
3.4 Apache	31
3.5 MySQL	32
3.6 PHP	32

4	ΜΕΤΑ-ΑΝΑΛΥΣΗ	33
5	ΜΕΘΟΔΟΣ ΚΑΙ ΔΙΑΔΙΚΑΣΙΑ.....	38
5.1	Αναζήτηση Πρωτογενών Δεδομένων	38
5.1.1	Gene Expression Omnibus (GEO).....	39
5.1.2	Πειραματικός Σχεδιασμός των Επιλεγμένων Σειρών.....	41
5.2	Επεξεργασία με Χρήση RStudio.....	45
5.2.1	Agilent Platform	45
5.2.2	CEA - DSV/DRR/SGF: Human 25k oligo v1.0 (GenePix).....	50
5.3	Περιγραφή της Βάσης Microarray.....	53
5.4	Υπολογισμός P-value (Bayesian Analysis).....	62
5.5	Μετα-ανάλυση (STATA)	68
5.6	Βιβλιογραφική Αναζήτηση (Quertle-Cosmic)	69
6	ΑΠΟΤΕΛΕΣΜΑΤΑ.....	70
6.1	Αρχικός Έλεγχος με Διαγράμματα Venn	70
6.2	Μετα-ανάλυση.....	71
6.3	WebGestalt	97
6.4	Απεικόνιση Βιολογικών Δικτύων (STRING).....	110
6.5	Συνοπτική Παρουσίαση των Αποτελεσμάτων.....	113
6.6	Γονίδια από την Βιβλιογραφία	118
6.7	Τελικό Διάγραμμα Venn	122
7	ΣΥΖΗΤΗΣΗ.....	125
8	ΒΙΒΛΙΟΓΡΑΦΙΑ.....	131

1 ΙΟΝΤΙΖΟΥΣΕΣ ΑΚΤΙΝΟΒΟΛΙΕΣ

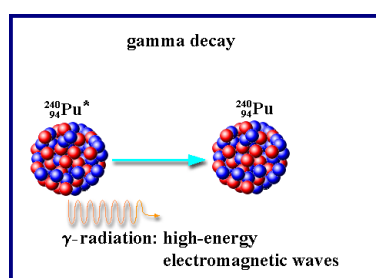
Ο ατομικός αριθμός ενός στοιχείου, δηλαδή ο αριθμός πρωτονίων του πυρήνα του, καθορίζει τις χημικές του ιδιότητες. Πυρήνες με ίδιο ατομικό αριθμό αλλά διαφορετικό μαζικό (άθροισμα πρωτονίων και νετρονίων) ονομάζονται ισότοπα. Τα περισσότερα στοιχεία στην φύση υπάρχουν ως μείγμα ισωτόπων. Οι φυσικές ραδιενεργές σειρές Θορίου, Ραδίου-Ουρανίου και Ακτινίου, περιλαμβάνουν τα στοιχεία που είναι από την φύση τους ραδιενεργά, τα οποία μπορούν και παράγουν με φυσικό τρόπο την ιοντίζουσα ακτινοβολία (UNEP, 2016). Ο όρος ιοντίζουσες/ιονίζουσες ακτινοβολίες (Ionizing Radiation IR) αναφέρεται στις ακτινοβολίες που μεταφέρουν την κατάλληλη ενέργεια ώστε κατά την αλληλεπίδραση τους με την ύλη να προκαλέσουν ιοντισμό των ατόμων της. Στις φυσικές πηγές ακτινοβολίας περιλαμβάνονται: η κοσμική που προέρχεται από το διάστημα, η ακτινοβολία εδάφους που οφείλεται στα ραδιενεργά ισότοπα των στοιχείων που περιέχονται στο έδαφος και σε ορισμένα πετρώματα (McAulay and Morgan, 1988) και τέλος η ραδιενέργεια στο ανθρώπινο σώμα λόγω του ισωτόπου του άνθρακα ^{14}C και του καλίου ^{40}K που περιέχεται στο βιολογικό ύδωρ, τα οποία και αποτελούν δομικά στοιχεία της οργανικής ύλης (IAEA, 2004). Στις τεχνητές πηγές ακτινοβολίας περιλαμβάνονται: η ακτινοβολία από την διάγνωση, την θεραπεία (ακτινολογία) αλλά και την χρήση ραδιοϊσοτόπων στην ιατρική (Little, 2003). Επίσης σε ευρύτερο πλαίσιο συνεισφέρουν τα ραδιενεργά απόβλητα και η ραδιενεργή σκόνη από πυρηνικές δοκιμές, ενώ οφείλει να σημειωθεί η επαγγελματική έκθεση από πυρηνικούς αντιδραστήρες και επιταχυντές. Σε γενικές γραμμές οι τεχνητές πηγές συνεισφέρουν σε μικρό ποσοστό έκθεσης του ανθρώπου στην ιοντίζουσα ακτινοβολία, ενώ πάνω από 90% αυτής οφείλεται σε φυσικές πηγές που διαδίδονται μέσω του αέρα, των τροφών αλλά και του ίδιου του ανθρώπινου σώματος.

1.1 Εκπομπή & Ιδιότητες Ιοντίζουσών Ακτινοβολιών

Κατά την αλληλεπίδρασή τους με την ύλη, οι δέσμες ιοντίζουσας ακτινοβολίας υπόκεινται σε απώλεια ενέργειας και εκτροπή από την αρχική διεύθυνση. Τα χαρακτηριστικά και τα αποτελέσματα κάθε αλληλεπίδρασης καθορίζονται κυρίως από το είδος της ακτινοβολίας (UNEP, 2016). Σε αυτές περιλαμβάνονται:

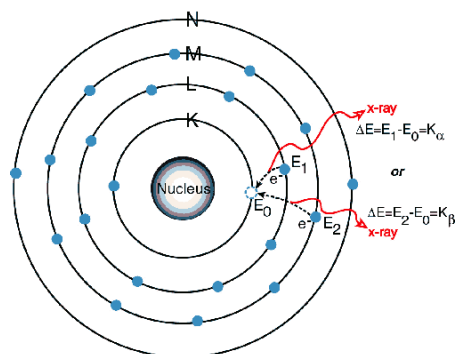
- Ηλεκτρομαγνητικές ακτινοβολίες: ακτίνες-Χ και ακτίνες-γ
- Σωματιδιακές ακτινοβολίες:
 - Σωματίδια-α (πυρήνες He)
 - Σωματίδια-β (ηλεκτρόνια και ποζιτρόνια)
 - Νετρόνια, νετρίνα (χαμηλής και υψηλής ενέργειας ουδέτερα σωματίδια)
 - Πρωτόνια ή άλλα βαρέα ιόντα (π.χ. ιόντα άνθρακα)

Ακτίνες-γ: πρόκειται για ακτινοβολία φωτονίων υψηλής ενέργειας, στην περιοχή μήκους κύματος των 10^{-1} έως 10^{-6} nm (περιοχή συχνοτήτων $5 \cdot 10^{19}$ ως $3 \cdot 10^{22}$ Hz) του ηλεκτρομαγνητικού φάσματος, που εκπέμπονται από ασταθείς πυρήνες. Ο μητρικός πυρήνας που παράγει το φωτόνιο δεν αλλάζει ως προς την μάζα ή τον ατομικό αριθμό παρά μόνο ελαττώνεται η ενέργειά του (Εικόνα 1). Καθώς το φωτόνιο δεν διαθέτει μάζα και φορτίο, η ακτινοβολία-γ είναι ιδιαίτερα διεισδυτική και μπορεί να ταξιδέψει μεγάλες αποστάσεις στον αέρα χάνοντας το μισό της ενέργειάς της περίπου κάθε 150m. Η θωράκιση για την ακτινοβολία-γ επιτυγχάνεται με παχύ ή πυκνό στρώμα υλικού με μεγάλο ατομικό αριθμό, όπως για παράδειγμα ο μόλυβδος.



Εικόνα 1: Παραγωγή ακτίνας-γ μέσω αποδιέγερσης διεγερμένης κατάστασης πλουτωνίου (Pu), όπου ο μητρικός πυρήνας δεν αλλάζει ως προς την μάζα και τον ατομικό αριθμό.

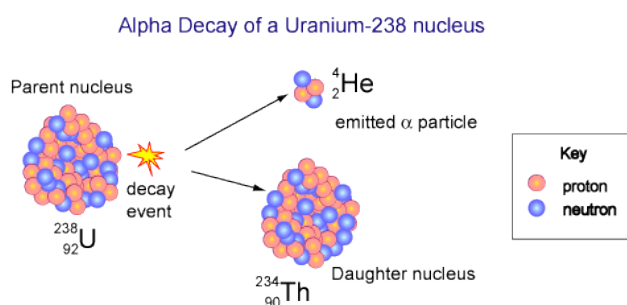
Ακτίνες-X: η ακτινοβολία αυτή έχει πολλά κοινά χαρακτηριστικά με την ακτινοβολία-γ με βασική διαφορά το ότι προέρχεται από ηλεκτρονιακό νέφος, ενώ συντίθεται από φωτόνια στην περιοχή μήκους κύματος των 1 έως 10^{-5} nm (περιοχή συχνοτήτων από $3 \cdot 10^{17}$ έως $5 \cdot 10^{19}$ Hz) του ηλεκτρομαγνητικού φάσματος. Πιο συγκεκριμένα παράγεται από ενεργειακές μεταβολές ηλεκτρονίων όπως είναι η μεταπήδηση από υψηλότερη σε χαμηλότερη ενεργειακή στάθμη μέσω εκπομπής (Εικόνα 2). Οι ακτίνες-X έχουν μεγαλύτερα μήκη κύματος και συνήθως μικρότερη ενέργεια από τις ακτίνες-γ. Παρότι μπορεί να προέλθουν από ορισμένα ραδιενεργά υλικά, συνήθως παράγονται από μηχανήματα που χρησιμοποιούνται στην ιατρική και την βιομηχανία.



Εικόνα 2: Μεταπήδηση διεγερμένου ηλεκτρονίου από τις στοιβάδες L και M στην K που συντελεί στην εκπομπή ακτίνων-X.

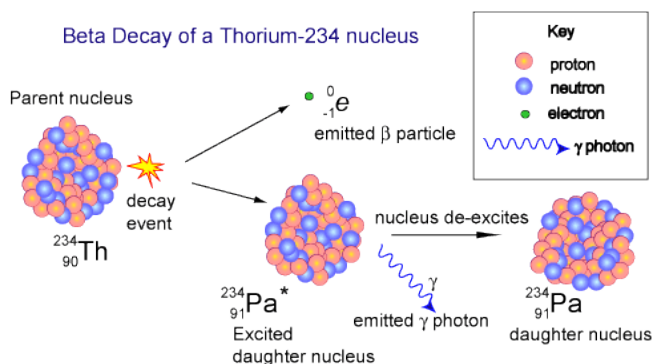
Στην περίπτωση ηλεκτρομαγνητικής ακτινοβολίας, στο υλικό-στόχο προκαλείται διέγερση και ιοντισμός των ατόμων ή μορίων. Οι κύριες αλληλεπιδράσεις που λαμβάνουν χώρα είναι: το φωτοηλεκτρικό φαινόμενο, το φαινόμενο *Compton* και η δίδυμη γένεση (pair production) αποτέλεσμα των οποίων είναι η παραγωγή ελευθέρων ηλεκτρονίων (Leroy and Rancoita, 2012). Μέσω των παραπάνω διαδικασιών η ενέργεια της δέσμης καταλήγει να απορροφάται πλήρως από τον στόχο. Αντίστοιχα κατά την αλληλεπίδραση ακτινοβολίας φορτισμένων σωματιδίων (όπως πρωτόνια ή σωματίδια-α), με την ύλη συμβαίνουν κυρίως ανελαστικές συγκρούσεις με τα ατομικά ηλεκτρόνια του στόχου και ελαστικές σκεδάσεις με τους πυρήνες των ατόμων του υλικού-στόχου, που έχουν ως αποτέλεσμα τον ιοντισμό και την διέγερση των ατόμων αλλά και τον ιοντισμό ή την διάσπαση των μορίων. Άλλα φαινόμενα και αιτίες απώλειας ενέργειας περιλαμβάνουν την εκπομπή δευτερευόντων ηλεκτρονίων από φαινόμενα απορρόφησης, την ακτινοβολία πέδησης (Low, 1958), την ακτινοβολία *Cherenkov* (Koch et al., 2006) και σπανιότερα ορισμένες πυρηνικές αντιδράσεις.

Ακτινοβολία-α: στην περίπτωση αυτή κάνουμε λόγο για σωματίδια που παράγονται από πυρηνική διάσπαση βαρέων, ασταθών πυρήνων, που δίνουν σωματίδιο που συνίσταται από δύο πρωτόνια και δύο νετρόνια, δηλαδή πυρήνα ηλίου ^4He . Κατά την διάσπαση αυτή ο μητρικός πυρήνας μεταπίπτει σε πυρήνα με ατομικό αριθμό κατά 2 μικρότερο από τον πρώτο (Εικόνα 3). Λόγω της μάζας και του φορτίου τους τα σωματίδια-α ταξιδεύουν στον αέρα μόνο λίγα εκατοστά και αλληλεπιδρούν έντονα με την ύλη. Ταξιδεύοντας μέσα στην ύλη χάνουν κινητική ενέργεια λόγω πολλαπλών συγκρούσεων με άλλα σωματίδια. Η διαδρομή τους μπορεί να εμποδιστεί από γυαλί, ένα φύλλο χαρτιού ή το δέρμα. Συνεπώς ο κίνδυνος που επιφυλάσσουν είναι μέσω της εισπνοής ή της κατάποσης από μολυσμένες επιφάνειες, καθώς έτσι μπορούν να φθάσουν απευθείας στο εσωτερικό του οργανισμού. Ως αποτέλεσμα των παραπάνω ένα σωματίδιο-α μπορεί να προκαλέσει μεγάλες βλάβες σε κύτταρα ή βιολογικούς ιστούς σε περίπτωση που εισέλθει στο σώμα.



Εικόνα 3: α-διάσπαση Ουρανίου που συνεπάγεται παραγωγή σωματιδίου-α και θυγατρικού πυρήνα ραδιενεργού θορίου.

Ακτινοβολία-β: Η β-διάσπαση συμβαίνει όταν ένα νετρόνιο είτε ελεύθερο είτε μέσα σε ένα ραδιοϊσότοπο αυθόρμητα διασπάται για να παράγει ένα πρωτόνιο, ένα ηλεκτρόνιο και ένα αντινετρίνιο ηλεκτρονίου (Εικόνα 4). Αντίστοιχα μπορεί να παραχθεί ένα πρωτόνιο, ένα ποζιτρόνιο και ένα νετρίνιο ηλεκτρονίου. Το εκπεμπόμενο ηλεκτρόνιο/ποζιτρόνιο έχει υψηλή κινητική ενέργεια και καλείται σωματίδιο-β. Η β-διάσπαση ανήκει στις ασθενείς πυρηνικές αλληλεπιδράσεις με φορέα αλληλεπίδρασης το μποζόνιο W^- . Λόγω της μικρής του μάζας μπορεί να ταξιδέψει μεγαλύτερες αποστάσεις μέσα στον αέρα (μερικά μέτρα), αλλά σταματά εύκολα στην ύλη μετά από διαδρομή λίγων εκατοστών. Συνεπώς έχει την δυνατότητα να διαπεράσει το δέρμα αλλά βασικό χαρακτηριστικό της ακτινοβολίας αυτής είναι ο κίνδυνος από κατάποση ραδιενεργής ουσίας μέσω εσωτερικής εκπομπής. Για την θωράκιση αυτής χρησιμοποιείται το γυαλί, το πλαστικό ή κάποιο ελαφρύ μέταλλο.



Εικόνα 4: β-διάσπαση Θωρίου που συνεπάγεται παραγωγή σωματιδίου-β (ηλεκτρονίου) και διεγερμένου θυγατρικού πυρήνα που αποδιεγείρεται με εκπομπή φωτονίου.

Ακτινοβολία νετρονίων: συνίσταται από ελεύθερα νετρόνια που εκπέμπονται ως αποτέλεσμα αυθόρμητης ή προκαλούμενης πυρηνικής σχάσης, ενώ οι ιδιότητες της καθορίζονται από το είδος των νετρονίων (γρήγορα / μεσαία / θερμικά). Έχει την δυνατότητα να ταξιδεύει έως και χιλιάδες μέτρα στον αέρα, μπορεί όμως να εμποδιστεί από υλικά πλούσια σε υδρογόνο, όπως το νερό και το μπετόν. Ο ιοντισμός που προκαλεί δεν είναι άμεσος, παρά μέσω της απορρόφησης των νετρονίων από σταθερό άτομο. Έτσι το άτομο καθίσταται ασταθές με αποτέλεσμα να αυξάνεται η πιθανότητα εκπομπής ιονίζουσας ακτινοβολίας διαφορετικού τύπου. Η ακτινοβολία νετρονίων αποτελεί τον μοναδικό τύπο ακτινοβολίας που έχει την δυνατότητα να καθιστά άλλα υλικά ραδιενεργά. Γενικότερα οι ακτινοβολίες ουδέτερων σωματιδίων προκαλούν μεταστοιχείωση των πυρήνων και καταστροφή των μορίων του υλικού στόχου.

1.2 Χαρακτηριστικά της Ακτινοβολίας

Ο ρυθμός απώλειας ενέργειας ως προς την διαδρομή που διανύει μια δέσμη φορτισμένων σωματιδίων στο υλικό δίνεται από την εξίσωση *Bethe-Bloch* (Εξίσωση 1) που περιλαμβάνει παραμέτρους της ακτινοβολίας αλλά και του υλικού του στόχου. Η ανασχετική ισχύς (*stopping power*), όπως επίσης αποκαλείται, εξαρτάται κυρίως από τον λόγο Z/A , που αποτελεί τον λόγο του ατομικού αριθμού προς τον μαζικό. Καθορίζεται από το υλικό του στόχου-απορροφητή και περιγράφει την διαδρομή της ακτινοβολίας στην ύλη (Ziegler et al., 2010).

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

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.

Επίσης για την απεικόνιση της εμβέλειας της ακτινοβολίας, η οποία καθορίζεται από τον ρυθμό με τον οποίο χάνει ενέργεια η δέσμη, λαμβάνονται υπόψιν κβαντικά και σχετικιστικά φαινόμενα. Για χαμηλές ενέργειες ακτινοβολίας απαιτούνται κυριότερα διορθώσεις φλοιού, ενώ για μεγάλες απαιτούνται διορθώσεις πυκνότητας. Για μικρές ορμές η κλίση της ανασχετικής ισχύος (dE/dx) ισούται με $1/\beta^2$ όπου β η σχετικιστική σταθερά που συνδέεται με την ορμή (Εξίσωση 2). Από το σημείο ελάχιστου ιοντισμού μέχρι να ξεκινήσει η πόλωση του μέσου, εμφανίζεται λογαριθμική αύξηση της ανασχετικής ισχύος, ενώ για μεγάλες ορμές ($\beta\cdot\gamma > 1000$) εμφανίζεται μικρή αύξηση, όπου και κυριαρχούν φαινόμενα *bremsstrahlung* (Γράφημα 1) (Porter, 1985).

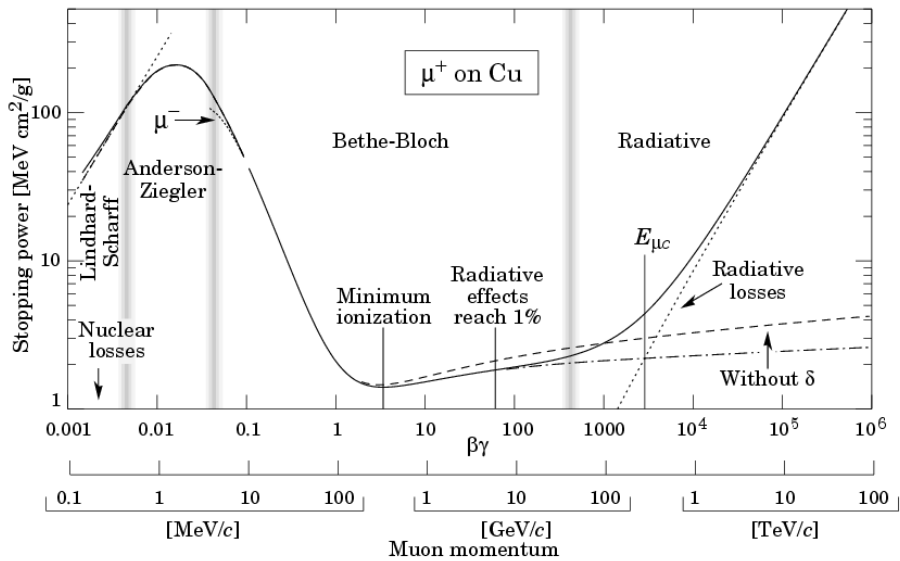
$$p = \beta\gamma mc$$

Εξίσωση 2: Εξίσωση σχετικιστικής ορμής με β και γ σχετικιστικές σταθερές, m η μάζα του σωματιδίου και c η ταχύτητα του φωτός.

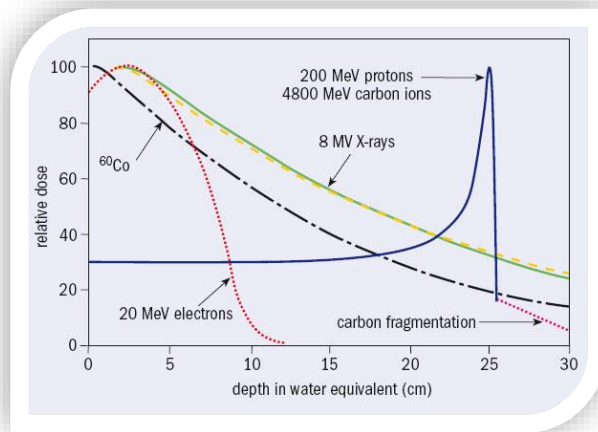
Η ακτινοβολία *Bremsstrahlung* πρόκειται για ηλεκτρομαγνητική ακτινοβολία, που παράγεται κατά την επιβράδυνση φορτισμένου σωματιδίου, λόγω κρούσης με φορτισμένο σωματίδιο μεγαλύτερης μάζας (Low, 1958). Η κινητική ενέργεια που χάνει το πρώτο (κινούμενο) μετατρέπεται σε φωτόνιο ώστε να ικανοποιείται ο νόμος διατήρησης της ενέργειας. Για μεγάλες ορμές, άρα και μεγάλες ενέργειες τα παραγόμενα φωτόνια (ακτινοβολία *bremsstrahlung*) εμφανίζουν μεγάλες συχνότητες, όπου και το φαινόμενο κορυφώνεται.

Πιο συγκεκριμένα στην περίπτωση μεγαλύτερων σωματιδίων (πρωτόνια, σωματίδια- α , βαρέα ιόντα) η απώλεια ενέργειας χαρακτηρίζεται από την καμπύλη *Bragg*

(Γράφημα 2), η οποία σηματοδοτεί μέγιστη εναπόθεση ενέργειας στο τέλος της διαδρομής που ακολουθείται από την ολική απορρόφηση της δέσμης (Kjellberg et al., 1983) και σχετίζεται με το φαινόμενο στραγγαλισμού (*straggling effect*).

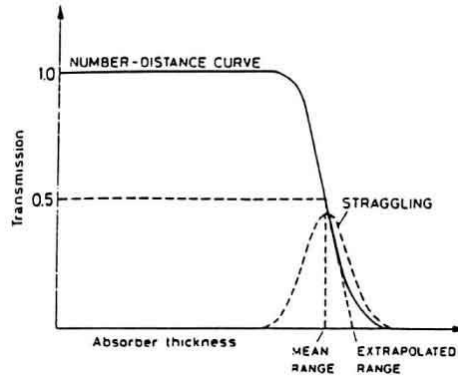


Γράφημα 1: Ανασχετική ισχύς μιονίων σε χαλκό σε σχέση με την σχετικιστική ορμή. Το φαινόμενο του στραγγαλισμού οφείλεται σε σωματίδια που συγκρατούνται σε συγκεκριμένο βάθος και προκαλούν ελάττωση της δέσμης (Γράφημα 3).



Γράφημα 2: Καμπύλη εναπόθεσης ενέργειας ως προς βάθος εισχώρησης διαφόρων ακτινοβολιών. Ξεχωρίζει η περίπτωση πρωτονίων και ιόντων άνθρακα όπου σχηματίζεται η χαρακτηριστική καμπύλη Bragg.

Καθώς η απώλεια ενέργειας είναι στατιστικής φύσεως, ορισμένα σωματίδια υπόκεινται σε λιγότερη ή περισσότερη απώλεια ενέργειας και έτσι η εμβέλεια καθίσταται μεγαλύτερη ή μικρότερη αντίστοιχα από την τυπική αναμενόμενη. Η μεταβολή αυτή στο ποσοστό συγκρούσεων-απώλειας προσεγγίζει Γκαουσιανή καμπύλη.



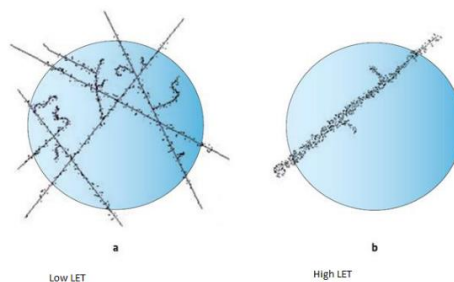
Γράφημα 3: Καμπύλη εκπομπής ως προς το πάχος του απορροφητή. Παρατηρείται το φαινόμενο του στραγγαλισμού, όπου *mean range*: 50% των σωματιδίων που εκπέμφθηκαν και *extrapolated range*: το πάχος όπου αναμένεται μηδενική εκπομπή-απορρόφηση.

Ως μέτρο για τον προσδιορισμό της κατανομής της ενέργειας στον ιστό χρησιμοποιείται η LET (*Linear Energy Transfer*) και υποδηλώνει την απώλεια ενέργειας της ακτινοβολίας dE λόγω συγκρούσεων με τα ηλεκτρόνια του στόχου για συγκεκριμένο μήκος διαδρομής dx στον στόχο (Εξίσωση 3). Αν και στο Διεθνές Σύστημα Μονάδων (S.I.) μετριέται σε (N) *newton*, στην περίπτωση της ιατρικής φυσικής και της ακτινοθεραπείας, συνήθως χρησιμοποιείται το keV/μm.

$$LET \equiv \frac{dE}{dx}$$

Εξίσωση 3: Γραμμική μεταβίβαση ενέργειας (*Linear Energy Transfer*).

Κατά την έκθεση σε ηλεκτρομαγνητικές ακτινοβολίες δημιουργούνται πολλαπλές τροχιές με μικρό ποσοστό ιονισμών ανά τροχιά και μικρότερη εναπόθεση ενέργειας συνολικά με αποτέλεσμα να χαρακτηρίζονται από χαμηλή τιμή LET (*low LET radiation*). Αντίθετα για τις σωματιδιακές ακτινοβολίες παρατηρείται εμφάνιση λίγων τροχιών με πυκνούς ιοντισμούς κατά μήκος μιας τροχιάς και αρκετά μεγαλύτερη εναπόθεση ενέργειας (Goodhead and Nikjoo, 1989) (Εικόνα 5). Έτσι χαρακτηρίζονται από υψηλή τιμή LET (*high LET radiation*) και συνολικά πιο στοχευμένη δράση (Held et al., 2016).



Εικόνα 5: (a) Τροχιά ιοντισμών low LET ακτινοβολίας και (b) Τροχιά ιοντισμών high LET ακτινοβολίας σε ένα κύτταρο.

Η ραδιενέργεια μιας πηγής μετριέται ως διασπάσεις ανά δευτερόλεπτο όπου 1Bq (*becquerel*) = 1 διάσπαση/sec ή εναλλακτικά με Ci (*curie*) όπου $1\text{Ci}=3.7 \cdot 10^{10} \text{Bq}$. Ως μέτρο της ραδιενέργειας που προσλαμβάνει ένας απορροφητής και κατ' επέκταση ένα έμβιο ον, χρησιμοποιούμε την απορροφούμενη δόση, η οποία αποτελεί το ποσό της ενέργειας που απορροφάται ανά μονάδα μάζας από ένα σύστημα και θεωρείται ο βέλτιστος τρόπος έκφρασης της απορροφούμενης ακτινοβολίας. Ως μονάδα αυτής χρησιμοποιείται κυρίως το Gy (*gray*) όπου $1\text{Gy}=1 \text{J/kg} = 100 \text{rad}$. Διακρίνουμε έτσι την ακτινοβολία σε υψηλές και χαμηλές δόσεις (*low και high dose*) θεωρώντας χαμηλές όσες είναι μικρότερες από 0.1Gy . Παρόλα αυτά στην παρούσα εργασία λάβαμε ως χαμηλές δόσεις μέχρι και 0.5Gy , όπως θεωρήθηκε στα πειραματικά δεδομένα που χρησιμοποιήθηκαν. Αντίστοιχα η έκθεση στην ακτινοβολία (ή δόση ιοντισμού) χαρακτηρίζει την δέσμη και αποτελεί μέτρο της ραδιενέργειας ως προς την ικανότητα να δημιουργήσει ιοντισμούς στον αέρα υπό κανονικές συνθήκες πίεσης και θερμοκρασίας. Καθώς μετριέται σε R (*roentgen*) όπου $1\text{R} = 2.58 \cdot 10^{-4} \text{C (coulomb)/kg αέρα}$ δεν μπορεί να χρησιμοποιηθεί απευθείας για την περιγραφή δόσης σε βιολογικό ιστό. Για τον σκοπό αυτό χρησιμοποιείται η ισοδύναμη δόση (*equivalent dose*), μονάδα της οποίας είναι το Sv (*sievert*), όπου $1\text{Sv}=100 \text{rad}=100 \text{Rem}$ και χρησιμοποιείται για την σύγκριση της «σχετικής βιολογικής αποτελεσματικότητας» RBE (*Relative Biological Effectiveness*) η οποία αποτελεί μια εμπειρική τιμή που προκύπτει από πειραματικές μετρήσεις και εξαρτάται άμεσα από την τιμή LET (Keith et al., 2012). Αντίστοιχα η ενεργός δόση (*effective dose*) χρησιμοποιείται για την εκτίμηση των κινδύνων έκθεσης του ανθρώπου στην ακτινοβολία και μετριέται επίσης σε Sv. Τέλος ως συλλογική δόση (*collective dose*) ορίζεται η δόση ανά άτομο σε Sv πολλαπλασιαζόμενη με τον αριθμό των ατόμων που εκτέθηκαν σε ακτινοβολία ανά χρόνο με μονάδες *man-sievert per year* (National Research Council, 2006). Χρησιμοποιείται συνήθως στην ακτινοπροστασία και σε υπολογισμούς πληθυσμιακής απόκρισης. Το ετήσιο ανεκτό όριο πρόσληψης ακτινοβολίας για τον οργανισμό που συσσωρεύεται και λειτουργεί αθροιστικά είναι 50mSv .

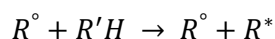
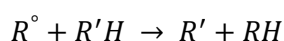
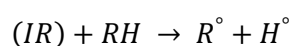
1.3 Το Χημικό Στάδιο Αλληλεπίδρασης

Οι βιολογικές συνέπειες της ιοντίζουσας ακτινοβολίας στον βιολογικό ιστό εξαρτώνται συνολικά από διάφορες παραμέτρους, όπως το είδος της ακτινοβολίας, η ποσότητά ακτινοβολίας που θα απορροφηθεί, η χωρική και χρονική κατανομή της ενέργειας που μεταφέρθηκε αλλά και το ίδιο το βιολογικό υλικό στόχο. Το φυσικό στάδιο αυτής της αλληλεπίδρασης σχετίζεται κυρίως με την μεταφορά ενέργειας και πραγματοποιείται σε χρόνο μικρότερο των 10^{-18}sec . Ακολουθεί το χημικό στάδιο όπου τα συστατικά των

κυττάρων αντιδρούν με ταχείες χημικές αντιδράσεις. Βασικό χαρακτηριστικό σε αυτό το στάδιο είναι η δημιουργία ελευθέρων ριζών.

- **Άμεσα και Έμμεσα Αποτελέσματα**

Κατά την αλληλεπίδραση των φορτισμένων σωματιδίων με την έμβια ύλη, λόγω της επαρκούς κινητικής ενέργειας που διαθέτουν προκαλούν άμεσο ιοντισμό. Μέσω του σπασίματος των χημικών δεσμών παράγονται ασύζευκτα ηλεκτρόνια υψηλής δραστηριότητας που οδηγούν στον σχηματισμό ελευθέρων ριζών (Εικόνα 6). Οι ελεύθερες ρίζες είναι άτομα ή μόρια ηλεκτρικά ουδέτερα, με ίσους αριθμούς πρωτονίων και ηλεκτρονίων, που η ηλεκτρονιακή τους κατανομή δίνει ένα ασύζευκτο ηλεκτρόνιο, με αποτέλεσμα να είναι εξαιρετικά χημικώς δραστικές.



Εικόνα 6: Αντιδράσεις παραγωγής ελευθέρων ριζών.

Στην περίπτωση ηλεκτρομαγνητικής ακτινοβολίας τα πράγματα είναι διαφορετικά. Καθώς ο πιο σημαντικός αποδέκτης ηλεκτρονίων στους βιολογικούς ιστούς είναι το μοριακό οξυγόνο έχουμε έμμεσο ιοντισμό μέσω της ραδιόλυσης του ύδατος (Εικόνα 7). Το οξυγόνο λόγω της δομής του, επιτρέπει την σύλληψη ελεύθερων ηλεκτρονίων σχηματίζοντας διάφορα «είδη» οξυγόνου που είναι υπεύθυνα για την δημιουργία χημικά ενεργών μορίων που περιέχουν οξυγόνο, τις ελεύθερες ρίζες (Riley, 1994).

- $(IR) + H_2O \rightarrow H_2O^+ + e^-$
- $H_2O^+ + H_2O \rightarrow H_3O^+ + HO^\circ$ ή εναλλακτικά
- $(IR) + H_2O \rightarrow H_2O^* \rightarrow H_2O + \gamma$
- ή $H_2O^* \rightarrow HO^\circ + H^\circ$ και έπειτα
- $e^- + H_2O \rightarrow H_2O^- + HO^- + H^\circ$

Εικόνα 7: Αντιδράσεις ραδιόλυσης του ύδατος.

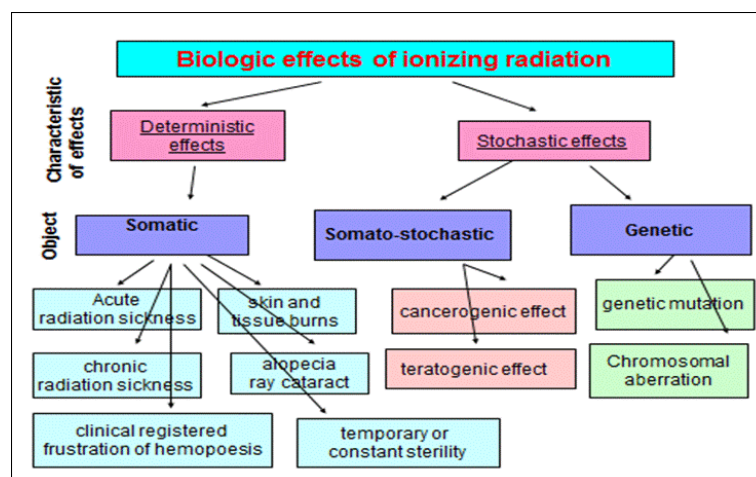
Οι ελεύθερες ρίζες που δημιουργούνται προκαλούν βλαβερές βιολογικές αντιδράσεις που στοχεύουν σε δομικά και λειτουργικά μόρια των κυττάρων όπως στα νουκλεϊκά οξέα, στις πρωτεΐνες και στα λιπίδια. Στην περίπτωση των πρωτεϊνών, το οξειδωτικό στρες προκαλεί αλλαγές στις ιδιότητες και την λειτουργία τους. Ειδικά στην περίπτωση πρωτεϊνών που προσδένονται στο DNA (*DNA-binding proteins*) μπορεί να προκληθούν αλλαγές στην γονιδιακή έκφραση, το οποίο οδηγεί σε μεταβολή συγκεντρώσεων πρωτεϊνών που σχετίζονται με τους μηχανισμούς επιδιόρθωσης ή την

εκκίνηση του μηχανισμού απόπτωσης (Wallace, 1998). Επίσης τα γεγονότα ιονισμού είναι πιθανό να επηρεάσουν απευθείας το γενετικό υλικό ιδιαίτερα στην περίπτωση του άμεσου ιονισμού, θέτοντας σε κίνδυνο την συνολική λειτουργία αυτού και του περιβάλλοντα χώρου.

Το Βιολογικό Στάδιο Αλληλεπίδρασης

1.3.1 Βιολογικές Αποκρίσεις

Οι βασικοί μηχανισμοί στους οποίους οφείλονται οι βλάβες σε χημικό επίπεδο είναι η παραγωγή ελευθέρων ριζών και η απελευθέρωση ηλεκτρονίων από τους μοριακούς δεσμούς, οι οποίες όμως καταλήγουν σε μεταβολή της δομής των βιομορίων του ιστού, με έμφαση στις πρωτεΐνες και το DNA. Τα αποτελέσματα από την έκθεση αρχικά κατηγοριοποιούνται με βάση το πόσο γρήγορα λαμβάνεται αυτή, σε οξέα (*acute*) και σε χρόνια (*chronic*). Τα οξέα προέρχονται από ακτινοβολήση σε όλο το σώμα, σε μικρό χρονικό διάστημα, με δόσεις των 100 mGy ή περισσότερο, ενώ τα αποτελέσματα τους είναι ικανά να εκδηλωθούν από λίγες ώρες μέχρι μερικές εβδομάδες μετά την έκθεση (*early effects*) και πολλά από αυτά προκαλούν την εμφάνιση συνδρόμων (Smith and Kao, 2004). Από την άλλη, τα χρόνια αποτελέσματα προέρχονται από σχετικά μικρές δόσεις ακτινοβολίας μέχρι 100 mGy, που απορροφούνται από το σώμα κατά παρατεταμένες χρονικές περιόδους.



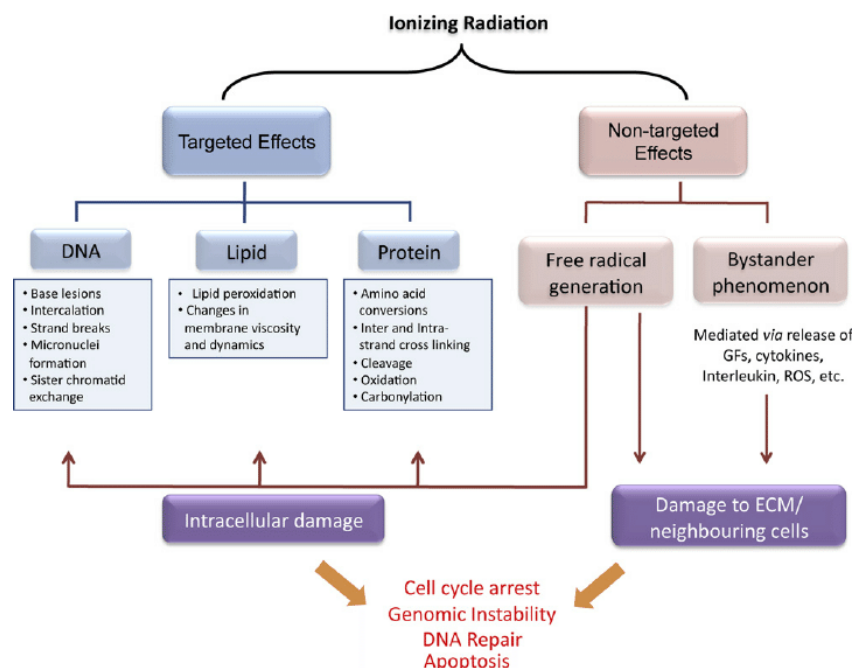
Εικόνα 8: Διαχωρισμός στοχαστικών και μη-στοχαστικών αποτελεσμάτων από έκθεση σε ιοντίζουσα ακτινοβολία.

Τα βιολογικά αποτελέσματα της έκθεσης σε ακτινοβολία χωρίζονται επίσης σε στοχαστικά (*stochastic*) και μη στοχαστικά (*non-stochastic*). Τα πρώτα χαρακτηρίζονται από βλάβες και μεταλλάξεις στα κύτταρα, ενώ τα δεύτερα καταλήγουν συνήθως σε κυτταρικό θάνατο, που οδηγεί σε υπερπληθυσμό νεκρών κυττάρων και αδυναμία αντικατάστασής τους λόγω έλλειψης χωρητικότητας (Hatzi et al., 2015) (Εικόνα 8). Τα στοχαστικά αποτελέσματα προκαλούνται έπειτα από έκθεση σε χαμηλή

δόση ακτινοβολίας (Valentin, 2005) και είναι πιο περίπλοκα ως προς τον καθορισμό τους σε αντίθεση με τα μη στοχευμένα, που μπορούν να διαγνωσθούν με ιατρική εξέταση λίγες εβδομάδες από την έκθεση και περιλαμβάνουν καψίματα του δέρματος και μείωση του ποσοστού των ερυθρών αιμοσφαιρίων.

Οι βιολογικές αποκρίσεις σε συνδυασμό με βιοφυσικές εκτιμήσεις συνθέτουν τα αποτελέσματα της έκθεσης της έμβιας ύλης στην ιονίζουσα ακτινοβολία. Η αλληλεπίδραση της ακτινοβολίας με οποιοδήποτε ζωντανό οργανισμό προκαλεί μία ποικιλία αποκρίσεων και ένα τοπικό ή γενικευμένο στρες εξαιτίας της παραγωγής ελεύθερων ριζών στην ακτινοβολημένη περιοχή ή γενικότερα στον οργανισμό μέσω συστημικών (μη στοχευμένων) επιδράσεων (Εικόνα 9), της απόκρισης στις βλάβες DNA (*DNA Damage Response*), των διαδικασιών επισκευής DNA και της έναρξη της φλεγμονώδους οδού (*pro-inflammatory pathway*) με αποτέλεσμα να επηρεάζεται άμεσα το ανοσοποιητικό σύστημα (Georgakilas et al., 2015).

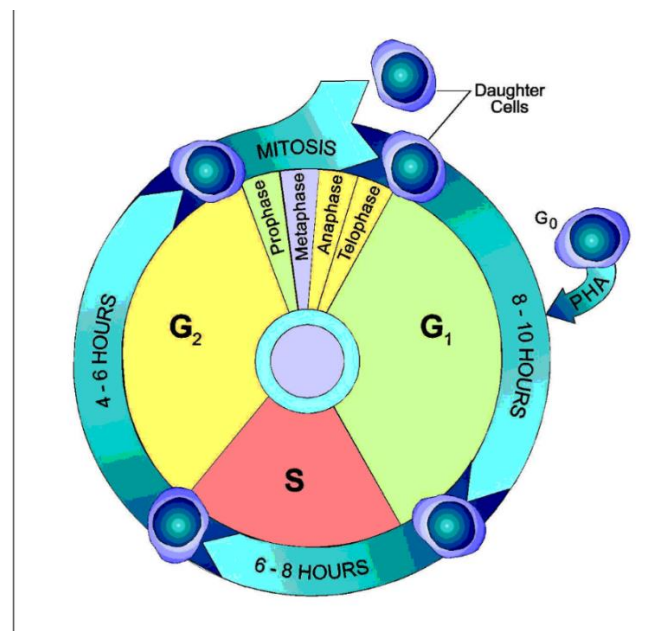
Επιπλέον βλάβες μπορεί να συμβούν στα μιτοχόνδρια και την κυτταρική μεμβράνη ενός κυττάρου με αποτέλεσμα να επηρεαστεί η συνολική λειτουργία του. Τέλος ορισμένοι μεταγραφικοί παράγοντες φαίνεται να ενεργοποιούνται λίγα λεπτά μετά την έκθεση σε κλινικής φύσεως ιονίζουσα ακτινοβολία (<4Gy). Σε αυτούς περιλαμβάνονται οι p53 και NF-κB. Ο καθένας από αυτούς επηρεάζει με διαφορετικό τρόπο την λειτουργία του κυττάρου ενώ το σύνολο των γονιδίων των οποίων η έκφραση αναστέλλεται ή καταστέλλονται από αυτούς είναι τεράστιο.



Εικόνα 9: Διαχωρισμός στοχευμένων και μη στοχευμένων επιδράσεων της ιονίζουσας ακτινοβολίας.

1.3.2 Δομή DNA και Κυτταρικός Κύκλος

Η κλασική άποψη για την δομή δεξιόστροφης διπλής έλικας του DNA (*Deoxyribo-Nucleic Acid*) διατυπώθηκε πρώτη φορά από τους *Watson* και *Crick* το 1953. Οι αζωτούχες βάσεις κάθε κλώνου είναι κάθετες ως προς τον άξονα του μορίου και συνιστούν το εσωτερικό του μορίου όντας υδρόφοβες. Αντίστοιχα σχηματίζεται ένας σκελετός από επαναλαμβανόμενα μόρια φωσφορικής ομάδας στο εξωτερικό μέρος του μορίου, που είναι υδρόφιλος, γεγονός που δείχνει πως είναι κατασκευασμένο ώστε να προστατεύει την αλληλουχία των βάσεων από χημικές αλλοιώσεις από το περιβάλλον του κυττάρου. Βασικό στοιχείο δηλαδή του DNA αποτελεί η ακολουθία των τεσσάρων διαφορετικών νουκλεοτιδίων που ορίζονται από τις τέσσερις βάσεις, αδενίνη(A), γουανίνη(G), που συγκαταλέγονται στις πουρίνες και τις θυμίνη(T), κυτοσίνη(C), που ανήκουν στις πυραμιδίνες. Στην ακολουθία αυτή εμπεριέχεται η γενετική πληροφορία. Η έλικα δημιουργείται από τους δύο κλώνους με βάση τον κανόνα της συμπληρωματικότητας των αζωτούχων βάσεων η οποία επιβάλλει ζεύγος A-T με δύο δεσμούς υδρογόνου και G-C με τρεις. Τέλος στις διαδικασίες αντιγραφής, μεταγραφής και μετάφρασης συμμετέχουν μονόκλινα μόρια RNA όμοιας δομής με αυτή του κλώνου του DNA, όπου η θυμίνη αντικαθίσταται από την ουρακίλη (U) ενώ το ζάχαρο που συνθέτει τον εξωτερικό σκελετό μαζί με την φωσφορική ομάδα είναι η ριβόζη σε αντίθεση με την δεσοξυριβόζη του μορίου DNA (Sinden, 1994).



Εικόνα 10: Διάγραμμα αναπαράστασης του κυτταρικού κύκλου.

Ένα ακόμη σημαντικό χαρακτηριστικό της αλυσίδας είναι ότι οι κλώνοι τοποθετούνται αντιπαράλληλα. Πιο συγκεκριμένα κάθε κλώνος έχει τα άκρα 3' και 5' που διαφέρουν χημικά και βιολογικά, γεγονός που δίνει την ιδιότητα της πόλωσης σε κάθε κλώνο.

Τέλος σημαντικό ρόλο στην κατανόηση της επίδρασης των ιοντιζουσών ακτινοβολιών διαδραματίζουν οι διαφορετικές φάσεις του κυτταρικού κύκλου (Εικόνα 10). Η φάση όπου εμφανίζεται η μεγαλύτερη ευαισθησία σε μεταβολές είναι η G2 που ακολουθείται από τη μίτωση ενώ οι επιδιορθωτικοί μηχανισμοί της κυτταρικής βλάβης λειτουργούν σε όλες τις φάσεις του κυτταρικού κύκλου αλλά είναι αποδοτικότεροι στη φάση S.

1.3.3 Βλάβες DNA

Η βιολογική επίδραση των ακτινοβολιών σχετίζεται άμεσα και με τις αλλαγές στον γενετικό κώδικα, που προκύπτουν από μεταβολές της δομής του DNA (International Atomic Energy Agency (IAEA), 2010). Οι πιθανές βλάβες του DNA που προέρχονται από έκθεση σε ιοντίζουσα ακτινοβολία είναι:

- Σπάσιμο ενός κλώνου της αλυσίδας (*Single Strand Break-SSB*)
- Διπλό σπάσιμο της αλυσίδας, δηλαδή των δύο κλώνων (*Double Strand Break-DSB*)
- 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 (διάδεσμοι - *crosslinks*)

Η επιδιόρθωση του DNA αναφέρεται σε ένα σύνολο διαδικασιών μέσω των οποίων ένα κύτταρο αντιλαμβάνεται, αναγνωρίζει και επιδιορθώνει βλάβες στα μόρια DNA που κωδικοποιούν το γονιδίωμα του. Η ικανότητα αποκατάστασης του DNA σε ένα κύτταρο για την ακεραιότητα του γονιδιώματός του, όπως και η επιτυχημένη ενεργοποίηση μονοπατιών προγραμματισμένου κυτταρικού θανάτου (απόπτωσης) στις περιπτώσεις που είναι αναγκαίο, διαδραματίζουν κυρίαρχο ρόλο στην κανονική λειτουργία των κυττάρων και κατ' επέκταση όλου του οργανισμού (Torroni and Das, 2015). Διάφορα γονίδια που αποδεικνύεται ότι επηρεάζουν τη συνολική διάρκεια ζωής ενός οργανισμού, φαίνεται να εμπλέκονται στην αποκατάσταση και την προστασία από βλάβες στο DNA.

1.3.4 Χαμηλές Δόσεις Ακτινοβολίας

Οι χαμηλές δόσεις ακτινοβολίας επάγουν μηχανισμούς επιδιόρθωσης DNA, διακυτταρικές οξειδοαναγωγικές αντιδράσεις και μονοπάτια που σχετίζονται με σημεία ελέγχου του κυττάρου. Σε ορισμένες περιπτώσεις ένα κύτταρο μπορεί να οδηγηθεί σε απόπτωση ή σε κυτταρικό θάνατο που σχετίζεται με την διαδικασία της μίτωσης. Αυτά συμβαίνουν ως απόπειρα προστασίας του κυττάρου και κατ' επέκταση του οργανισμού

από την έκθεση σε ιοντίζουσα ακτινοβολία. Όμως οι χαμηλές δόσεις φαίνεται να προκαλούν αβέβαια αποτελέσματα, που δεν ακολουθούν απαραίτητα κάποιο γραμμικό μοντέλο, σε αντίθεση με τις υψηλές και για αυτό αποτρέπεται ο ακριβής προσδιορισμός των μηχανισμών απόκρισης (Lagarde, 2003). Η μελέτη της επίδρασης των χαμηλών δόσεων υποθέτει επίσης πως το κύτταρο αποτελεί μονάδα εκτίμησης του κινδύνου για τον οργανισμό αγνοώντας τα μη στοχευμένα αποτελέσματα της ακτινοβολίας που συνδέονται με το *bystander effect* (Εικόνα 9). Ο όρος *bystander effect* στην ραδιοβιολογία, αναφέρεται σε ένα φαινόμενο κατά το οποίο τα μη ακτινοβολημένα κύτταρα εμφανίζουν μια απόκριση παρόμοια με τα «γειτονικά» τους ακτινοβολημένα κύτταρα, πιθανότερα λόγω παραγωγής χημικών σημάτων και σημάτων ανασοαπόκρισης από τα δεύτερα (Georgakilas, 2015) και ως απόδοση στα ελληνικά μπορεί να χρησιμοποιηθεί ο όρος «φαινόμενο του παρατυχόντος». Κοινή παράμετρος των παραπάνω φαινομένων είναι η δράση και η μετάδοση του τοπικού στρες λόγω ακτινοβολίας στον κυτταρικό πληθυσμό ή τον οργανισμό σε απόσταση που μπορεί να κυμαίνεται από λίγα χιλιοστά έως και αρκετά εκατοστά. Τα σήματα αυτά μεταδίδονται μέσω διακυτταρικών χασμοσυνδέσμων (Dauer et al., 2010). Τα αποτελέσματα αυτών περιλαμβάνουν: βλάβες DNA, χρωμοσωμικές ανωμαλίες και γονιδιωματική αστάθεια, τροποποιημένη γονιδιακή έκφραση, απόπτωση, μεταβολές στη μεταγωγή σημάτων, ραδιενεργό απόκριση και νεοπλαστικό μετασχηματισμό (Kadhim et al., 2013). Η κατανόηση των μηχανισμών πίσω από αυτό το φαινόμενο βασίζεται στη συμμετοχή των οξειδωτικών και φλεγμονώδων αποκρίσεων που προκαλούν.

Οι βλάβες DNA που συμβαίνουν, σε συνδυασμό με πολλαπλές μοριακές και επιγενετικές μεταβολές μπορεί να έχουν ως αποτέλεσμα την δημιουργία καρκινικού όγκου. Από βιολογική σκοπιά, θεωρείται ότι υπάρχει ένα κατώφλι για τα μη στοχαστικά αποτελέσματα καθώς αυτά συμβαίνουν όταν έχουν υπάρξει βλάβες σε μεγάλο αριθμό κυττάρων σε συγκεκριμένο κρίσιμο χρονικό διάστημα και τα οποία το σώμα δεν έχει τη δυνατότητα να διορθώσει ή αντικαταστήσει. Αν τα κύτταρα αυτά συνεχίσουν κανονικά να διαιρούνται αντί να οδηγηθούν σε απόπτωση οι βλάβες διαιωνίζονται. Ένα τέτοιο λάθος μπορεί να οδηγήσει σε καρκινογένεση (Auvinen et al., 1998), για το λόγο αυτό και θεωρείται κατά βάση μονοκυτταρική ως προς την προέλευσή της.

Ως καρκίνο χαρακτηρίζουμε μη φυσιολογικά κύτταρα που διαιρούνται ανεξέλεγκτα και μπορούν να εισβάλλουν σε κοντινούς ιστούς. Τα καρκινικά κύτταρα μπορούν επίσης να εξαπλωθούν σε άλλα μέρη του σώματος μέσω του αίματος και των λεμφικών συστημάτων. Υπάρχουν αρκετοί κύριοι τύποι καρκίνου ανάλογα με το σημείο του σώματος από το οποίο προήλθε (μονοκυτταρική προέλευση), καθώς ο καρκίνος

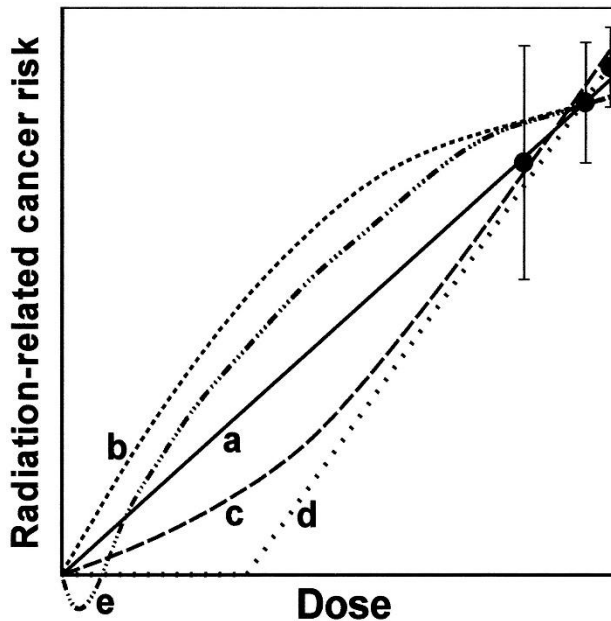
σχηματίζεται από μεταλλαξογόνο βλάβη σε ένα κύτταρο και κυριότερα στο γενετικό υλικό του πυρήνα. Παρόλα αυτά η κακοήθεια (όγκος) που σχηματίζεται, αποτελεί αποτέλεσμα μιας ολόκληρης σειράς κυτταρικών αλλαγών σε γονίδια που μπορεί να συμβαίνουν για μεγάλο χρονικό διάστημα μέσω της ανάπτυξης όλων των συστατικών των κυττάρων και που θα καταλήξει στη διαίρεση για δημιουργία δύο θυγατρικών κυττάρων, τα οποία θα πρέπει να συνεχίσουν να διαιρούνται για τον σχηματισμό του όγκου. Φαίνεται ότι ο κίνδυνος εμφάνισης καρκίνου είναι υπαρκτός στην περίπτωση έκθεσης σε χαμηλές δόσεις (Little et al., 2009) αν και κάποτε θεωρούνταν αβλαβείς. Η διαδικασία συσχέτισης της δόσης ακτινοβολίας και τον συγκεκριμένων μηχανισμών που ενεργοποιούνται όμως είναι δύσκολη εξαιτίας της έλλειψης αξιόπιστων βιοδεικτών (Martin et al., 2014).

Σχετικά με την απόκριση σε χαμηλές δόσεις λοιπόν, αν και είναι γνωστό ότι υπάρχει μεγάλος αριθμός ποσοτικών προσδιορισμών μετάλλαξης, δεν είμαστε σε θέση να ανιχνεύσουμε τον τρόπο και την συχνότητα των παραγόντων που συντελούν στις βλάβες και τις κυτταρικές δυσλειτουργίες. Η δόση της ακτινοβολίας που χορηγείται σε έναν οργανισμό επιτρέπει εκτιμήσεις του αριθμού των τροχιών ιονισμού σε μεμονωμένα κύτταρα. Σε πολύ χαμηλές δόσεις πολύ λίγα κύτταρα θα υποστούν άμεσο χτύπημα, καθιστώντας τα μη ικανά να ενεργοποιήσουν την κατάλληλη απόκριση για έκθεση σε ακτινοβολία. Το γεγονός αυτό, ειδικά σε δόσεις χαμηλότερες του 1 mGy, έχει ως αποτέλεσμα να μην δεχτούν όλα τα κύτταρα κάποια εναπόθεση ενέργειας (Sykes and Day, 2007). Έτσι το κύτταρο που εκτίθεται σε πολύ χαμηλή δόση ακτινοβολίας, κάποιες φορές αποτυγχάνει στην επιτυχή ενεργοποίηση των μηχανισμών επιδιόρθωσης DNA κυρίως μέσω της ATM. Η πρωτεϊνική κινάση σερίνης/θρεονίνης ATM, ενεργοποιείται όταν συμβαίνουν DSBs, φωσφορυλιώνοντας αρκετές πρωτεΐνες, όπως η p53, η H2AX και άλλους ογκο-κατασταλείς, που με την σειρά τους είναι υπεύθυνες για την ενεργοποίηση σημείων ελέγχου του κυτταρικού κύκλου, πολλαπλών επιδιορθωτικών μηχανισμών αλλά και μηχανισμών απόπτωσης. Φαίνεται πως η ποσότητα της φωσφορυλιωμένης H2AX (γ -H2AX) σε ινοβλάστες (fibroblasts) που ακτινοβολήθηκαν με χαμηλή δόση είχε διατηρηθεί στα φυσιολογικά επίπεδα, ενώ η ποσότητα του μεταγραφικού παράγοντα NF- κ B που ρυθμίζει τα επίπεδα ελευθέρων ριζών είχε αυξηθεί (Dauer et al., 2010). Καθώς η έκθεση σε χαμηλές δόσεις προκαλεί την παραγωγή ελευθέρων ριζών κατά μεγαλύτερο ποσοστό από τις θραύσεις διπλής έλικας (DSBs), ο NF- κ B αποτελεί πιθανότερο βιοδείκτη για τα αποτελέσματα αυτής.

Η φλεγμονώδης απόκριση όπως και η επαγωγή του αναστολέα της απόπτωσης (XIAP), καθορίζονται από τον μεταγραφικό παράγοντα NF-κB. Κατά την θεραπεία με ακτινοβολία χαμηλής δόσης προκαλούνται αντιφλεγμονώδη αποτελέσματα μεταβάλλοντας τον φαινότυπο των ανοσοκυττάρων. Όταν τα κύτταρα της ανοσοαπόκρισης (μακροφάγα) ενεργοποιηθούν, αποτελούν κύρια πηγή αρκετών κυτοκινών και χημοκινών, που διαδραματίζουν σημαντικό ρόλο στην ρύθμιση της ανοσοαπόκρισης. Φαίνεται ότι ειδικά η κυτοκίνη IL-1β έχει πολύπλευρη επίδραση στο ανοσοποιητικό σύστημα δρώντας σε πολλαπλά ανοσοκύτταρα. Έχει παρατηρηθεί σημαντικά μειωμένη ποσότητα της βιολογικώς δραστηκής μορφής (διασπασμένη μορφή) της IL-1β μετά από θεραπεία με χαμηλές δόσεις ακτίνων-Χ, οι οποίες ρυθμίζουν τον φαινότυπο των ανοσοποιητικών κυττάρων με έναν τρόπο εξαρτώμενο από τον NF-κB. Επιπλέον τα επίπεδα της ογκο-κατασταλτικής πρωτεΐνης p35 καθώς και η ενεργή διασπασμένη μορφή της IL-1β (p17) μειώνονται εντός των ακτινοβολημένων μακροφάγων (Lodermann et al., 2012). Συνεπώς η τροποποιημένη έκφραση και έκκριση της IL-1β ρυθμίζεται μέσω του NF-κB μετά από έκθεση των ενεργοποιημένων μακροφάγων σε χαμηλές δόσεις ακτίνων-Χ.

Τέλος αξίζει να σημειωθεί πως μεμονωμένη έκθεση σε χαμηλή δόση ακτινοβολίας, προκαλεί διαδικασίες που μπορούν να μειώσουν τον κίνδυνο καρκινογένεσης, καθώς πολλαπλές επιδράσεις προσαρμογής (*adaptive effects*) μεταδίδονται σε γειτονικά μη στοχευμένα κύτταρα και τα προστατεύουν από έκθεση σε high LET ακτινοβολία (de Toledo et al., 2011). Τα κύτταρα αυτά είναι ικανότερα στο να αποκαταστήσουν σωστά βλάβες DNA και επομένως λιγότερο πιθανό να μετασχηματιστούν στο νεοπλαστικό φαινότυπο.

Οι κίνδυνοι από έκθεση σε χαμηλές δόσεις σε σύγκριση με έκθεση σε υψηλές λοιπόν μπορεί να θεωρούνταν κάποτε αμελητέοι, αλλά τα νεότερα δεδομένα συσχετίζουν την έκθεση σε χαμηλές δόσεις με την δημιουργία όγκων (Εικόνα 11), ενώ παράλληλα υπάρχουν πολλά ζητήματα να διευθετηθούν ως προς: την πιθανότητα εμφάνισης καρκίνου έπειτα από την έκθεση, τους ακριβείς μηχανισμούς απόκρισης και πως είναι δυνατό να τους διαχειριστούμε καταλλήλως σε κάθε περίπτωση (Schuler and Gutzwiller, 1991). Για αυτό το λόγο απαιτούνται ευρείες επιδημιολογικές μελέτες για τον ακριβή προσδιορισμό των επιδράσεων αυτού του είδους ακτινοβολίας (Brenner et al., 2003).



Εικόνα 11: Σχηματική αναπαράσταση διαφόρων πιθανών σεναρίων για τους κινδύνους ακτινοβολίας έπειτα από έκθεση σε πολύ χαμηλές δόσεις, οι οποίες μπορεί να είναι συμβατές με επιδημιολογικά δεδομένα από έκθεση σε υψηλότερες δόσεις ακτινοβολίας. Όλες οι σχέσεις δόσης-απόκρισης αποτελούν πιθανές περιγραφές ογκογένεσης (Brenner et al., 2003).

1.3.5 Υψηλές Δόσεις Ακτινοβολίας

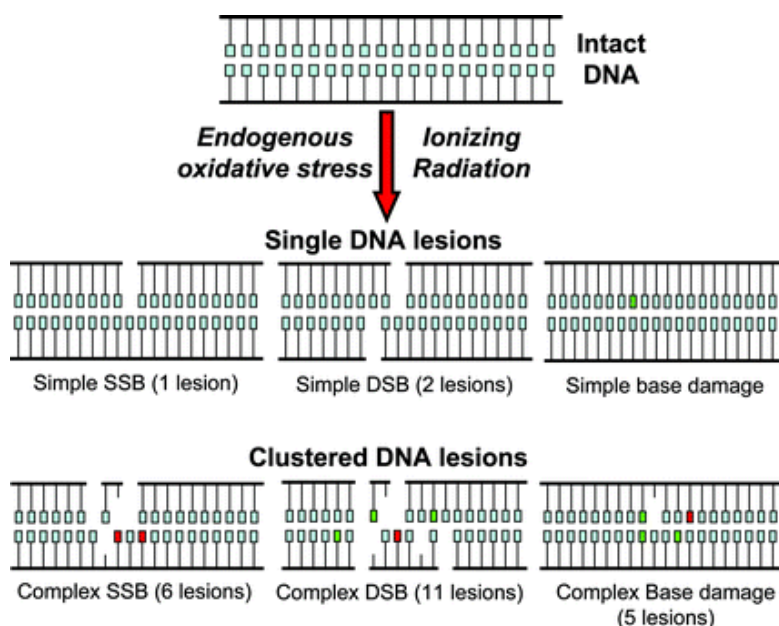
Οι βιολογικές επιδράσεις έπειτα από έκθεση σε υψηλά επίπεδα ακτινοβολίας σκιαγραφούνται αρκετά ξεκάθαρα σε σχέση με την έκθεση σε χαμηλές δόσεις καθώς το μοντέλο που τις περιγράφει είναι γραμμικό (Dauer et al., 2010). Τα μοριακά φαινόμενα που παρατηρούνται είναι κοινά για χαμηλές και υψηλές δόσεις, αλλά η αύξηση της δόσης συντελεί στην αύξηση της πιθανότητας προκαλούμενων ιοντισμών. Έτσι σκιαγραφείται με σιγουριά η βιολογική απόκριση, που σχετίζεται με τις διαδικασίες επιδιόρθωσης DNA, τον οξειδωτικό μεταβολισμό και τα σημεία ελέγχου του κυτταρικού κύκλου (*cell cycle checkpoints*) λόγω της διαφορικής ρύθμισης διαδικασιών σηματοδότησης. Το αποτέλεσμα της έκθεσης σε πολύ υψηλή δόση ακτινοβολίας χαρακτηρίζεται ως οξεία ακτινοπληξία, ενώ συχνή συνέπεια αποτελεί και η δημιουργία καρκίνου. Τα συμπτώματα οφείλονται κυρίως σε μη στοχαστικά αποτελέσματα και περιλαμβάνουν κοκκίνισμα του δέρματος, πόνο, έλκος, διάρροια ακόμη και αναιμία, αν ακτινοβοληθεί σημαντική ποσότητα μυελού των οστών. Στην περίπτωση που τα αναπαραγωγικά κύτταρα (γαμέτες) έρθουν σε επαφή με την ακτινοβολία είναι πιθανό να εμφανιστούν μεταλλάξεις στους απογόνους ή εναλλακτικά να προκληθεί στειρώση.

1.3.6 Ο ρόλος του LET–Ομαδοποιημένες Βλάβες

Η πολυπλοκότητα και η σοβαρότητα μίας σύνθετης βλάβης αυξάνεται με την αύξηση της LET και κατά συνέπεια η μεταλλαξιογόνος και η καρκινογόνος δράση. Στην περίπτωση high LET ακτινοβολίας, σε αντίθεση με τις προσαρμοστικές αποκρίσεις σε κύτταρα/ιστούς έπειτα από έκθεση σε χαμηλή δόση και παράλληλα low LET ακτινοβολία, εμφανίζονται επιδράσεις στρες που επιμένουν. Το οξειδωτικό και θερμικό

στρες δεν περιορίζεται στον ιστό-στόχο αλλά διαδίδεται και σε γειτονικά κύτταρα/ιστούς. Οι σύνθετες βλάβες DNA αποτελούν την υπογραφή της high LET ακτινοβολίας. Μία βλάβη χαρακτηρίζεται ως σύνθετη όταν αποτελείται από βλάβες που έχουν μικρή απόσταση και σχηματίζουν ένα σύμπλεγμα βλαβών στο γενετικό υλικό (Georgakilas et al., 2013). Οι δύο βασικές κατηγορίες σύνθετων βλαβών DNA είναι οι δίκλωνες θραύσεις (DSBs) και οι μη-δίκλωνες οξειδωτικές.

Οι ομαδοποιημένες βλάβες DNA θεωρούνται ότι αποτελούν κακώς επισκευασμένες βλάβες που καταλήγουν σε κυτταροτοξικές και μεταλλαξιογόνες δράσεις και χρωμοσωμική αστάθεια. Παρόλο που πολλαπλά μονοπάτια για την επαγωγή και διαιώνιση της γονιδιωματικής αστάθειας είναι πιθανά, θεωρείται ότι οι DSBs και άλλες μορφές ομαδοποιημένων βλαβών εμπλέκονται στη γονιδιωματική αστάθεια και την γέννηση χρωμοσωμικών ανωμαλιών. Οι χημικές διαδικασίες και οι κυτταρικοί μηχανισμοί επιδιόρθωσης σύνθετων βλαβών στο DNA δεν έχουν χαρακτηριστεί πλήρως σε αντίθεση με την περίπτωση των μεμονωμένων βλαβών όπως μια οξείδωση βάσης ή το σπάσιμο μιας αλυσίδας.



Εικόνα 12: Εξειδανικευμένη αναπαράσταση διάφορων τύπων ομαδοποιημένων βλαβών, συμπεριλαμβανομένων σύνθετων θραυσμάτων μονής έλικας (SSB) και διπλών θραυσμάτων (DSBs) που μπορούν να επέλθουν έπειτα από έκθεση σε ιοντίζουσα ακτινοβολία (Georgakilas, 2008).

Η low LET ακτινοβολία μπορεί να δημιουργήσει συστάδες (*clusters*) με έως και 10 αλλοιώσεις ανά σύμπλεγμα, ενώ η high LET ακτινοβολία δύναται να προκαλέσει ζημιά ακόμη μεγαλύτερης πολυπλοκότητας μέχρι και 25 αλλοιώσεις ανά σύμπλεγμα (Εικόνα 12). Η καλύτερη γνώση περί της επαγωγής και επιδιόρθωσης σύνθετων βλαβών είναι απαραίτητη για τον υπολογισμό των παραγόντων κινδύνου ακτινοβολίας high LET

σχετικά με την πρόκληση καρκίνου, ιδιαίτερα στις περιπτώσεις που χρησιμοποιείται για θεραπευτικό σκοπό (Hada and Georgakilas, 2008).

Οι ομαδοποιημένες βλάβες DNA σε αντίθετους κλώνους λοιπόν, παρουσιάζουν μία αξιοσημείωτη πολυπλοκότητα που τις καθιστά δύσκολο να επιδιορθωθούν και ικανές να προκαλέσουν μεταλλαξογόνες ή ακόμη και θανατηφόρες επιπτώσεις σε ένα κύτταρο σε σχέση με τυχόν απομονωμένες (Georgakilas et al., 2004). Πιο συγκεκριμένα κατά την επιδιόρθωση σημείων όπου δημιουργούνται συστάδες βλαβών (*heavily clustered sites*) το κύτταρο οδηγείται σε μεγάλες διαγραφές αλληλουχιών γεγονός το οποίο συντελεί σε γέννηση περίπλοκων μεταλλάξεων. Αποτέλεσμα της έκθεσης στην περίπτωση της high LET ακτινοβολίας είναι ο σχηματισμός καινοφανών αλληλουχιών σε ακτινοβολημένες περιοχές του κυττάρου που συμμετέχουν σε σύνθετες αναδιατάξεις και καταλήγουν σε ορατές χρωμοσωμικές αλλοιώσεις που συνήθως συνιστούν στη δημιουργία καρκινογενέσεων εάν δεν επέλθει κυτταρικός θάνατος (Singleton et al., 2002; Stewart et al., 2011).

2 ΒΙΟΠΛΗΡΟΦΟΡΙΚΕΣ ΜΕΘΟΔΟΙ ΕΠΕΞΕΡΓΑΣΙΑΣ

Ως βιοπληροφορική ορίζουμε την επιστήμη που ασχολείται με την συλλογή, την ταξινόμηση, την αποθήκευση και την ανάλυση βιοχημικών και βιολογικών πληροφοριών μέσω χρήσης ηλεκτρονικών υπολογιστών, όπως εφαρμόζονται κυρίως στη μοριακή γενετική και τη γονιδιωματική. Χρησιμοποιεί διάφορες στατιστικές μεθόδους και γενικότερα εφαρμοσμένα μαθηματικά, που σε συνδυασμό με την επιστήμη της βιολογίας στοχεύει στη προσέγγιση βιολογικών προβλημάτων και επιδιώκει τη διατύπωση νέων αρχών. Πιο συγκεκριμένα έχει σκοπό την οργάνωση, την διαχείριση και την κατανόηση της πληροφορίας που σχετίζεται με βιομακρομόρια όπως το DNA, το RNA, τις πρωτεΐνες και τους πολυσακχαρίτες. Στους στόχους της βιοπληροφορικής ανάλυσης περιλαμβάνονται: η οργάνωση των βιολογικών δεδομένων και η εύκολη πρόσβαση σε αυτά, καθώς και η συσσώρευση νέων δεδομένων με αποδοτικό και εύχρηστο τρόπο, που περιλαμβάνει ανάπτυξη μεθόδων και υπολογιστικών εργαλείων για την εξαγωγή των επιθυμητών πληροφοριών από αυτά. Τελικό στάδιο της ανάλυσης αποτελεί η χρήση των εργαλείων αυτών για την περιγραφή και ερμηνεία των δεδομένων με ένα βιολογικά αποδεκτό τρόπο.

2.1 Μικροσυστοιχίες

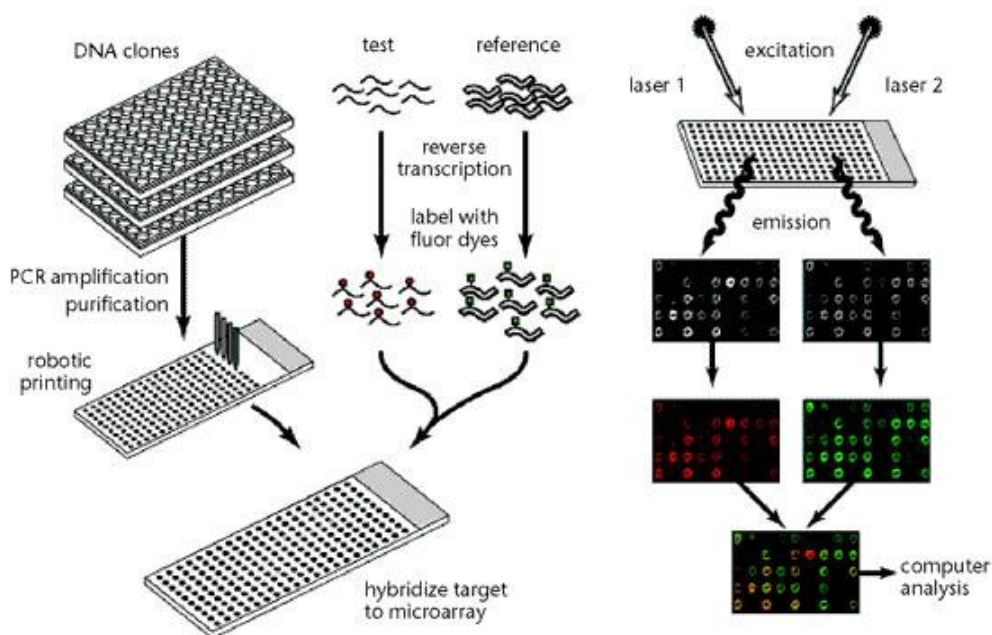
Οι μικροσυστοιχίες (*microarrays*) επιτρέπουν την ανάλυση της γονιδιακής έκφρασης, της ποικιλότητας, της αλληλουχίας του DNA και των επιπέδων και τροποποιήσεων των πρωτεϊνών με μαζική και παράλληλη επεξεργασία. Αποτελούν μια τεχνολογία με ευρείες εφαρμογές σε τομείς όπως η γονιδιωματική, η πρωτεωμική και η διαγνωστική καθώς δίνουν τη δυνατότητα ανάλυσης ολόκληρου του μεταγραφώματος ενός οργανισμού σε ένα μόνο πείραμα (Lander et al., 2001). Μέσω της διατύπωσης του βιολογικού ερωτήματος, τον σχεδιασμό του εκάστοτε πειράματος και την προετοιμασία των δειγμάτων γίνεται η κατάλληλη επεξεργασία των δεδομένων για την εξαγωγή του συμπεράσματος. Η γονιδιακή έκφραση είναι άμεσα συσχετιζόμενη με τις βιολογικές λειτουργίες και οι μικροσυστοιχίες υπόσχονται τεράστιο όγκο δεδομένων πάνω σε ανθρώπινες ασθένειες, στη φαρμακευτική και ορμονική δράση, σε διανοητικές ασθένειες, στο μεταβολισμό και σε αρκετά ακόμα κλινικά θέματα (Schena, 1996). Υπάρχουν τρεις βασικές κατηγορίες μικροσυστοιχιών:

- Μικροσυστοιχίες πρωτεϊνών
- Μικροσυστοιχίες ιστών
- Μικροσυστοιχίες cDNA

Στην παρούσα εργασία θα ασχοληθούμε με την τρίτη κατηγορία, τις μικροσυστοιχίες γενετικού υλικού. Τα προτερήματα αυτού του τύπου μικροσυστοιχιών προκύπτουν από την εξαιρετική εξειδίκευση και δυνατότητα συσχέτισης που επιτυγχάνουν λόγω της συμπληρωματικότητας των βάσεων. Ένα αντίγραφο DNA ενός απομονωμένου γονιδίου αποτελεί σχεδόν τέλειο αντιδραστήριο για ειδική και ποσοτική ανίχνευση όπως και για εύρεση αλληλουχιών (Duggan et al., 1999). Θεωρητικά κάθε ανιχνευτής (probe) αντιστοιχεί σε ένα γονίδιο (Heller, 2002). Οι μικροσυστοιχίες DNA λοιπόν είναι μια διάταξη μεγάλου αριθμού ανιχνευτών DNA που αντιπροσωπεύουν συγκεκριμένους γενετικούς τόπους. Συστοιχίες από χιλιάδες διακριτές αλληλουχίες DNA τυπώνονται σε πλάκες μικροσκοπίου από γυαλί με την χρήση ρομποτικού συστοιχίτη (*arrayer*). Οι ανιχνευτές ακινητοποιούνται σε στερεή επιφάνεια με τεχνικές σύγχρονης νανοτεχνολογίας. Σε κάθε probe υπάρχει μια «ομάδα» μονόκλωνου DNA που σχετίζονται με μια συγκεκριμένη διεργασία (Lou et al., 2001). Έτσι επιτρέπεται η ανάλυση της γονιδιακής έκφρασης και της ποικιλότητας με μαζική επεξεργασία. 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, σε υψηλή θερμοκρασία ή έπειτα από έκθεση σε ιοντίζουσα ακτινοβολία (Brown and Botstein, 1999).

Για την ετοιμασία του βιολογικού υλικού του δείγματος, εκχυλίζεται το ολικό RNA από τον ιστό. Μέσω αντίστροφης μεταγραφής του mRNA παράγουμε συμπληρωματικό DNA (cDNA) το οποίο σημαίνεται με φθορίζουσα χρωστική ή διαφορετικές χρωστικές για κάθε δείγμα στην περίπτωση δικάναλων μικροσυστοιχιών (Duggan et al., 1999). Πιο συγκεκριμένα εκμεταλλευόμαστε το polyA που εντοπίζεται μετά την 3'UTR περιοχή του mRNA στα ευκαριωτικά κύτταρα και έτσι με προσθήκη oligo-dT (ολιγονουκλεοτίδια περίπου 20 βάσεων με θυμίνη), αντίστροφης μεταγραφάσης και φθορίζοντων μονοδεσοξυριβονουκλεοτιδίων (A, T, G, C), συντίθεται το σημασμένο cDNA. Στη συνέχεια επιστρώνεται στην μικροσυστοιχία ώστε να υβριδοποιηθεί. Απαιτείται χημικό «ξέπλυμα» του πλακιδίου ώστε να διατηρηθούν μόνο τα πλήρως υβριδοποιημένα τμήματα της αλυσίδας. Σε αυτό το στάδιο, λόγω της συμπληρωματικότητας των βάσεων τα φθορίζοντα μονομερή που αντιστοιχούν σε κάθε ανιχνευτή εμφανίζονται σε διαφορετικές αναλογίες (Heller, 2002). Το επόμενο βήμα είναι η σάρωση της επιφάνειας της μικροσυστοιχίας, απ' όπου προκύπτει μια ψηφιακή εικόνα, που προέρχεται από τη διέγερση των μορίων σήμανσης που βρίσκονται στους στόχους και φθορίζουν σε συγκεκριμένα μήκη κύματος, φανερώνοντας τις συγκεντρώσεις σε κάθε

ανιχνευτή. Κατ' αυτό τον τρόπο, αυξημένη φωτεινότητα σημαίνει αυξημένη έκφραση γονιδίου στο δείγμα (Schena, 2002).



Εικόνα 13: Απεικόνιση της διαδικασίας κατασκευής και ανάγνωσης πειράματος μικροσυστοιχιών για την μελέτη της γονιδιακής έκφρασης (Duggan et al., 1999).

Σε ένα πείραμα ανάλυσης γονιδιακού προφίλ δύο χρωμάτων (δικαναλικές μικροσυστοιχιές) όπου οι στόχοι από τα δύο διαφορετικά δείγματα βάφονται ξεχωριστά εξάγεται μια διαφορετική εικόνα μετά την σάρωση λόγω διαφορετικού μήκους κύματος που εκπέμπει η φθορίζουσα ουσία. Συνήθως το ένα δείγμα έχει σημειωθεί με κόκκινο φθορίζον μόριο (Cy5) και το άλλο με πράσινο (Cy3) (Εικόνα 13). Οι στόχοι που βρίσκονται σε περίσσεια από το κάθε δείγμα για κάθε γονίδιο, θα υπερισχύσουν έναντι των άλλων και θα καταλάβουν περισσότερους ανιχνευτές (Dudoit et al., 2002). Θα δούμε κόκκινο τον ανιχνευτή αν το συγκεκριμένο γονίδιο υπερεκφράζεται στα κύτταρα του πρώτου δείγματος, θα δούμε πράσινο τον ανιχνευτή αν ισχύει το αντίστοιχο για το δεύτερο δείγμα, θα δούμε κίτρινο τον ανιχνευτή αν η έκφραση είναι παρόμοια και, τέλος, θα δούμε μαύρο τον ανιχνευτή αν όταν δεν υπάρχει καθόλου έκφραση (Cheung et al., 1999). Η εικόνα αυτή και οι τιμές έκφρασης προκύπτουν από ειδικούς σαρωτές. Στη συνέχεια προχωράμε σε ανάλυση των δεδομένων αυτών, από την οποία προκύπτουν ποσοτικοποιημένα δεδομένα για τις εκφράσεις (Butte, 2002; Lyng et al., 2004). Τα δεδομένα αυτά επεξεργάζονται στον ηλεκτρονικό υπολογιστή με πληθώρα αλγορίθμων, ώστε να απαλειφθούν τα σφάλματα και να δώσουν συμπεράσματα τα οποία ο άνθρωπος δε θα μπορούσε να εξάγει, λόγω του μεγάλου τους όγκου.

Για την απάντηση του βιολογικού ερωτήματος σε αυτή την ανάλυση, επιλέχθηκαν πειράματα που περιλάμβαναν υγιή δείγματα από ανθρώπινο ιστό. Σε κάθε μελέτη

υπάρχουν μη ακτινοβολημένα (*control samples*) και ακτινοβολημένα δείγματα με χαμηλή και κάποιες φορές υψηλή δόση εάν περιλαμβάνονταν στην μελέτη. Απώτερος σκοπός στην μελέτη αυτή είναι η διαφορική έκφραση των γονιδίων μεταξύ ακτινοβολημένων και μη δειγμάτων με έμφαση στην έκθεση σε χαμηλή δόση ιοντίζουσας ακτινοβολίας.

2.2 Πακέτο limma

Το άνοιγμα και η επεξεργασία των δεδομένων των μικροσυστοιχιών στην παρούσα εργασία έγινε με χρήση της προγραμματιστικής γλώσσας R στο ολοκληρωμένο περιβάλλον ανάπτυξης (*Integrated Development Environment*) RStudio (RStudio, 2012). Η R αποτελεί ένα περιβάλλον λογισμικού για στατιστικούς υπολογισμούς και γραφικά. Μία μεγάλη συλλογή πακέτων της R που χρησιμοποιείται ευρέως, ειδικά στην βιοπληροφορική είναι το Bioconductor (Gentleman et al., 2004), το οποίο θα χρησιμοποιηθεί σε αυτή την μελέτη και αποτελεί ελεύθερο λογισμικό. Βασικές διεργασίες για την ανάλυση των δεδομένων αποτελεί η διόρθωση σήματος υποβάθρου και η κατάλληλη κανονικοποίηση. Κάθε πλατφόρμα εμφανίζει μικρές διαφορές ως προς την διαχείριση της μέσω της R, το βασικό πακέτο όμως που χρησιμοποιείται σε όλες είναι το limma (Linear Models for Microarray and RNA-seq Data). Το πακέτο limma αποτελεί πακέτο λογισμικού για την ανάλυση πειραμάτων γονιδιακής έκφρασης, κυρίως με χρήση γραμμικών μοντέλων, με βασικό στόχο την αξιολόγηση της διαφορικής έκφρασης (Ritchie et al., 2015). Οι μέθοδοι ανάλυσης που περιλαμβάνει μπορούν να εφαρμοστούν σε διαφορετικές τεχνολογίες στις οποίες περιλαμβάνονται οι μικροσυστοιχίες, η RNA-seq, η ποσοτική PCR και πολλές τεχνολογίες πρωτεϊνών (Smyth et al., 2002).

2.2.1 Η Μέθοδος “normexp” και η Τιμή offset

Για την διόρθωση υποβάθρου επιλέχθηκε η μέθοδος normexp (Ritchie et al., 2007). Θεωρώντας μικροσυστοιχία δύο καναλιών χρωμάτων *Red* και *Green* υπολογίζεται η συνολική έκφραση για το ένα χρώμα (Εξίσωση 4). Έτσι έπειτα από μετασχηματισμούς βασισμένους στην Γκαουσιανή κατανομή των τιμών S, B τυπώνονται δύο λίστες από τιμές που εκφράζουν την έκφραση-φωτεινότητα (Εξίσωση 5). Σε περιπτώσεις μονοκαναλικών μικροσυστοιχιών το ένα χρώμα παραλείπεται στις εξισώσεις.

$$R_f = R_b + B + S$$

Εξίσωση 4: Υπολογιζόμενη έκφραση για ένα χρώμα, σε ένα spot μέσου της μεθόδου *normexp*. Όπου R_f η ένταση έκφρασης στο προσκήνιο (*foreground*), R_b η ένταση στο υπόβαθρο (*background*), S η φωτεινότητα που αντιστοιχεί στην ένταση της πραγματικής έκφρασης και ακολουθεί εκθετική κατανομή ενώ B είναι η ένταση υποβάθρου που δεν συμπεριλαμβάνεται στο R_b , η οποία ακολουθεί εκθετική κατανομή.

$$R = \text{normexp}(R_f - R_b) + \text{offset}$$

$$G = \text{normexp}(G_f - G_b) + \text{offset}$$

Εξίσωση 5: Τελικές τιμές έκφρασης για τα δύο χρώματα έπειτα από διόρθωση υποβάθρου.

Σχετικά με την τιμή *offset*, επειδή κατά την κανονικοποίηση οι τιμές θα λογαριθμιστούν (Εξίσωση 6), μας ενδιαφέρει στον λόγο έκφρασης των δύο χρωμάτων να μην προκύψει αρνητικός λογάριθμος ($R/G < 1$) ως τελική κανονικοποιημένη τιμή έκφρασης, οπότε συνίσταται $\text{offset} = 25$ (Smyth et al., 2002). Στην περίπτωση των μονοχρωματικών μικροσυστοιχιών διατηρείται μόνο ο αριθμητής του κλάσματος του λογαρίθμου. Συνεπώς για ένα χρώμα θέσαμε την τιμή $\text{offset} = 1$ για να αποφύγουμε πιθανές αρνητικές τιμές ($R < 1$) σε αντίθεση με την τιμή $\text{offset} = 0$ (Silver et al., 2009)

$$M = \log_2 \left(\frac{\text{normexp}(R_f - R_b) + \text{offset}}{\text{normexp}(G_f - G_b) + \text{offset}} \right)$$

Εξίσωση 6: Τελικές κανονικοποιημένες τιμές έκφρασης για διχρωματική μικροσυστοιχία.

2.2.2 Κανονικοποίηση Ποσοστημορίου

Η ανάγκη για κανονικοποίηση των τιμών των εντάσεων έκφρασης προκύπτει λόγω πολλαπλών ανεπιθύμητων παρεκκλίσεων που προκύπτουν από τεχνικά «λάθη». Αυτές οι παρεκκλίσεις οφείλονται σε μεταβολές των συνθηκών του εκάστοτε πειράματος που δεν μπορούν να αποφευχθούν και που σε συνδυασμό με τα βιολογικά αποτελέσματα μπορεί να οδηγήσουν σε ψευδείς ανακαλύψεις. Η κανονικοποίηση ποσοστημορίου (*quantile normalization*), την οποία και θα χρησιμοποιήσουμε, αναπτύχθηκε εξ αρχής και στοχευμένα για μικροσυστοιχίες γονιδιακής έκφρασης (Bolstad et al., 2003). Η μέθοδος αυτή θεωρεί ότι η στατιστική κατανομή κάθε δείγματος είναι η ίδια. Η μέση κατανομή, η οποία λαμβάνεται από τον μέσο όρο έκφρασης των γονιδίων σε κάθε δείγμα, χρησιμοποιείται ως αναφορά έτσι ώστε όλες οι κατανομές να είναι συγκρίσιμες μεταξύ τους.

Πίνακας 1: Υποθετική κατανομή για την εφαρμογή κανονικοποίησης ποσοστημορίου.

Genes	Control Sample	Low Dose Sample	High Dose Sample
A	5	7	1
B	4	6	5
C	6	4	4
D	7	5	3
E	3	3	7

Πίνακας 2: Τελική ταξινόμηση για κάθε δείγμα με εφαρμογή της μεθόδου κανονικοποίησης ποσοστημορίου.

Genes	Control Sample	Low Dose Sample	High Dose Sample
A	iii (4.7)	v (7)	i (2.3)
B	ii (3.7)	iv (5.7)	iv (5.7)
C	iv (5.7)	ii (3.7)	iii (4.7)
D	v (7)	iii (4.7)	ii (3.7)
E	i (2.3)	i (2.3)	v (7)

Πιο συγκεκριμένα για δύο ή περισσότερες κατανομές, οι κατανομές ταξινομούνται κατά αύξουσα σειρά, και μετά ορίζεται ο μέσος όρος, συνήθως ο αριθμητικός, των κατανομών (Εξίσωση 7). Έτσι η υψηλότερη τιμή σε όλες τις περιπτώσεις γίνεται ο μέσος όρος των υψηλότερων τιμών, η δεύτερη υψηλότερη τιμή γίνεται ο μέσος όρος των τιμών της δεύτερη υψηλότερης τιμής κοκ (Πίνακας 1, Πίνακας 2).

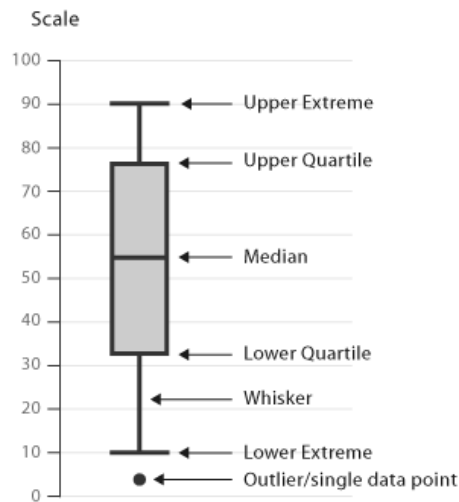
$$i = \frac{i_c + i_l + i_h}{N_i} = 2.3$$

Εξίσωση 7: Μέσος όρος έκφρασης από όλα τα δείγματα έπειτα από ταξινόμηση.

Έτσι δημιουργούνται πανομοιότυπες κατανομές ως προς τις στατιστικές τους ιδιότητες. Αυτόματα με την εντολή αυτή από το πακέτο `limma`, οι κανονικοποιημένες τιμές λογαριθμίζονται με βάση το 2 καθώς υπακούουν στον νόμο του *Benford* ή νόμο του πρώτου ψηφίου (Formann, 2010). Ο νόμος αυτός προκύπτει από την φαινομενολογική παρατήρηση της κατανομής της συχνότητας των αρχικών ψηφίων σε πολλά φυσικά σύνολα αριθμητικών δεδομένων και υποδηλώνει ότι σε αυτά τα σύνολα τα μη σημαντικά ψηφία δυσανάλογα συχνά απαντώνται ως σημαντικά ψηφία. Η λογαρίθμιση λοιπόν είναι απαραίτητη για την εξάλειψη αυτού του φαινομένου (Qiu et al., 2013).

2.2.3 Έλεγχος με Χρήση Θηκογράμματος

Το θηκόγραμμα (*boxplot*) αποτελεί μια μέθοδο γραφικής απεικόνισης αριθμητικών δεδομένων στην παραστατική στατιστική με χρήση τεταρτημορίων (Εικόνα 14). Οι αποστάσεις μεταξύ των διαφορετικών τμημάτων των κουτιών απεικονίζουν την διασπορά και την ασυμμετρία των δεδομένων (Pagano, 1992) ενώ αναπαρίστανται και οι ακραίες τιμές ως τελείες (*outliers*). Το κουτί υποδηλώνει ότι το 50% των τιμών είναι μεγαλύτερες από την τιμή στη διάμεσο (*median*) ενώ 25% των συνολικών τιμών τοποθετείται πάνω από το άνω τεταρτημόριο και 25% κάτω από το κάτω τεταρτημόριο αντίστοιχα. Οι μύστακες καθορίζουν την μέγιστη και ελάχιστη τιμή εξαιρώντας τις ακραίες τιμές (Roeven, 1986).



Εικόνα 14: Σχηματική αναπαράσταση ενός θηκογράμματος.

3 ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

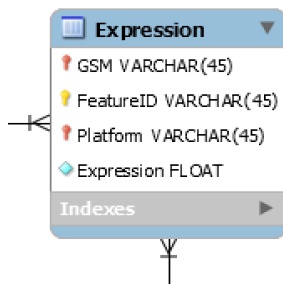
Ως βάση δεδομένων ορίζουμε μία συλλογή από συστηματικά οργανωμένα σχετιζόμενα δεδομένα που είναι ηλεκτρονικά κατατεθειμένα με σκοπό την εύκολη διαχείριση, την αποθήκευση και την επεξεργασία τους. Ο τρόπος που τα δεδομένα θα συνδεθούν διαδραματίζει κυρίαρχο ρόλο στην δημιουργία μιας βάσης δεδομένων. Γενικότερα τα δεδομένα θα πρέπει να έχουν λογική συνέχεια και νόημα καθώς μια βάση αντικατοπτρίζει ένα περιβάλλον του πραγματικού κόσμου. Κατά πλειοψηφία τα δεδομένα αναλύονται σε εγγραφές και πεδία. Μια εγγραφή αποτελεί μια συγκεκριμένη ομάδα πληροφοριών που εισάγεται στη βάση δεδομένων. Οι πληροφορίες αυτές μπορεί να αφορούν σε μια πρωτεΐνη, ένα γονίδιο ή ακόμα και ένα πρόσωπο. Κάθε μια από τις εγγραφές, αναλύεται σε πεδία που περιέχουν διάφορα στοιχεία απαραίτητα για την περιγραφή της π.χ. το όνομα της πρωτεΐνης. Τα πεδία είναι πάντοτε ίδια σε κάθε εγγραφή και ποτέ δεν υπάρχουν δυο ίδια πεδία σε μια εγγραφή.

3.1 Μοντέλο Διασύνδεσης Δεδομένων

Υπάρχουν τέσσερα μοντέλα βάσεων δεδομένων: Το ιεραρχικό, το δικτυακό, το σχεσιακό και το αντικειμενοστραφές. Συνήθως χρησιμοποιείται το σχεσιακό μοντέλο λόγω της απλής και κατανοητής δομής του. Αρχικά το μοντέλο αυτό προϋποθέτει την υποστήριξη της ανεξαρτησίας των δεδομένων ώστε να μην απαιτούνται αλλαγές στα προγράμματα της εφαρμογής στην περίπτωση που επιθυμούνται αλλαγές στην φυσική δομή και οργάνωση της βάσης. Επίσης αποφεύγεται η εμφάνιση/αποθήκευση δεδομένων κατά πολλαπλό τρόπο σε διαφορετικές περιοχές της βάσης δεδομένων και διατηρείται η ακεραιότητα και η συνέπεια των αυτών. Τέλος υποστηρίζεται η ανάπτυξη γλωσσών χειρισμού δεδομένων που διευκολύνουν τη διατύπωση ερωτημάτων προς το Σύστημα Διαχείρισης Βάσεων Δεδομένων.

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 (*data type*) και τη μορφοποίηση (*format*) (Darwen, 2014). 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 συνήθως επηρεάζει το χρόνο επεξεργασίας και επομένως λαμβάνεται υπόψη.



Εικόνα 15: Ο πίνακας με το όνομα Expression (από την βάση δεδομένων microarray που δημιουργήσαμε), έχει 4 γραμμές με ονομασίες: GSM, FeatureID, Platform και Expression. Οι γραμμές αυτές αποτελούν τα χαρακτηριστικά του πίνακα. Δεξιά από τα χαρακτηριστικά είναι ο τύπος των δεδομένων: VARCHAR, FLOAT και σε παρένθεση είναι το πεδίο ορισμού τους. Ο πίνακας έχει βαθμό 4.

Ένα χαρακτηριστικό που μπορεί να διακρίνει τις διάφορες στήλες του πίνακα, ονομάζεται πρωτεύον κλειδί. Για παράδειγμα τα πρωτεύοντα κλειδιά στον πίνακα Expression (Εικόνα 15) είναι τα GSM, Platform και FeatureID αλλά το Expression δεν μπορεί να θεωρηθεί πρωτεύον κλειδί, καθώς μπορεί να υπάρχουν περισσότερες από μία ίδιες εκφράσεις. Στην συγκεκριμένη περίπτωση το κλειδί καλείται σύνθετο καθώς απαιτούνται περισσότερα από ένα χαρακτηριστικά για να το συνθέσουν. 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: Είναι στενά συνδεδεμένη με την έννοια του κλειδιού μιας σχέσης. Κάθε στήλη πρέπει να προσδιορίζεται μοναδικά τουλάχιστον από το πρωτεύον κλειδί του πίνακα.
- Περιορισμοί αναφορών: 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 όπως στην παραπάνω περίπτωση του FeatureID. Η ακεραιότητα αναφορών επιβάλλει η τιμή ενός ξένου κλειδιού να έχει αντίστοιχη τιμή στον αρχικό πίνακα.

3.2 Συσχετίσεις

Οι πίνακες συνδέονται μέσω των εξής κατευθυνόμενων συσχετίσεων: μια εγγραφή ενός πίνακα συσχετίζεται με μία και μόνο μία εγγραφή ενός άλλου πίνακα (ένα προς ένα, 1:1), μία εγγραφή ενός πίνακα συσχετίζεται με πολλές εγγραφές ενός άλλου πίνακα αλλά και κάθε εγγραφή του δεύτερου πίνακα συσχετίζεται με μία και μόνο μία εγγραφή του πρώτου πίνακα (ένα προς πολλά, 1:N) και σε μία εγγραφή ενός πίνακα αντιστοιχούν πολλές εγγραφές ενός άλλου πίνακα και σε κάθε εγγραφή του δεύτερου πίνακα αντιστοιχούν πολλές εγγραφές του πρώτου πίνακα (πολλά προς πολλά, M:N). Υπάρχουν επίσης προαιρετικές προϋποθέσεις συμμετοχής των πινάκων (όπου ένας πίνακα δεν χρειάζεται να συσχετίζεται με τον άλλο πίνακα καθόλου). Τα πρωτογενή κλειδιά του πίνακα από τον οποίο ξεκινάει η συσχέτιση εξάγονται ως ξένα κλειδιά στον πίνακα που καταλήγει η συσχέτιση. Εάν τα ξένα κλειδιά γίνονται πρωτεύοντα κλειδιά στον καταλήγοντα πίνακα, η συσχέτιση ονομάζεται ταυτοποιούσα, ενώ εάν γίνονται δευτερεύοντα η συσχέτιση ονομάζεται μη ταυτοποιούσα (Darwen, 2014).

3.3 Διάγραμμα Συσχετίσεων Οντοτήτων (ERD)

Το Διάγραμμα Συσχετίσεων Οντοτήτων (Chen, 1976) είναι ένα μοντέλο δεδομένων για την περιγραφή μιας βάσης δεδομένων με έναν περιληπτικό τρόπο. Το σύστημα χαρακτήρων και συμβόλων που χρησιμοποιείται κατά κόρον από τα εργαλεία δημιουργίας Διάγραμμα Συσχετίσεων Οντοτήτων ονομάζεται *Crow's foot*. Οι πίνακες αντιπροσωπεύονται ως κουτιά και οι συσχετίσεις με γραμμές μεταξύ των κουτιών. Τα διαφορετικά σχήματα στα άκρα αυτών των γραμμών αντιπροσωπεύουν τους τύπους των συσχετίσεων. Οι συνεχείς γραμμές αναπαριστούν τις ταυτοποιούσες συσχετίσεις ενώ, οι διακεκομμένες γραμμές τις μη ταυτοποιούσες.

3.4 Apache

Για την δημιουργία της βάσης microarray χρησιμοποιήθηκε ο συνδυασμός */Apache/PHP/MySQL, που αποτελεί μία πλατφόρμα εκτέλεσης ιστοσελίδων. Γνωστή

και ως *AMP, όπου το πρώτο αρχικό αντιστοιχεί στην πλατφόρμα, στην οποία εγκαθίστανται ο Apache, η PHP και η MySQL (π.χ. Windows). Χρησιμοποιήθηκε για την εγκατάσταση αυτών σε Windows 10 το έτοιμο προς εγκατάσταση πακέτο XAMPP.

Ο Apache HTTP γνωστός και απλά σαν Apache είναι ένας εξυπηρετητής του παγκόσμιου ιστού (web). Όποτε ένας χρήστης επισκέπτεται ένα ιστότοπο το πρόγραμμα πλοήγησης (browser), επικοινωνεί με έναν διακομιστή (server) μέσω του πρωτοκόλλου HTTP, ο οποίος αποστέλλει το περιεχόμενο των ιστοσελίδων στο πρόγραμμα πλοήγησης. Ο Apache χρησιμοποιείται και σε τοπικά δίκτυα σαν διακομιστής συνεργαζόμενος με Συστήματα Διαχείρισης Βάσης Δεδομένων όπως η MySQL (Roberts et al., 2006).

3.5 MySQL

Η MySQL είναι ένα Σύστημα Διαχείρισης Σχεσιακών Βάσεων Δεδομένων που μετρά εκατομμύρια εγκαταστάσεις. Το πρόγραμμα τρέχει έναν εξυπηρετητή (server) παρέχοντας πρόσβαση πολλών χρηστών σε ένα σύνολο βάσεων δεδομένων. Ο κωδικός του εγχειρήματος είναι διαθέσιμος μέσω της GNU (General Public License), καθώς και μέσω ορισμένων ιδιόκτητων συμφωνιών. Ανήκει και χρηματοδοτείται από μία και μοναδική κερδοσκοπική εταιρία, τη σουηδική MySQL AB, η οποία σήμερα ανήκει στην Oracle. Η MySQL είναι δημοφιλής βάση δεδομένων για διαδικτυακά προγράμματα και ιστοσελίδες.

3.6 PHP

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 (Yu and Yi, 2010). 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 που παράγει κώδικα 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 Mac OS X υποστηρίζει εξ ορισμού την εκτέλεση κώδικα PHP.

4 ΜΕΤΑ-ΑΝΑΛΥΣΗ

Η μετα-ανάλυση αποτελεί ένα στατιστικό εργαλείο με σκοπό την επεξεργασία των δεδομένων και των αποτελεσμάτων που ερευνούν ένα συγκεκριμένο ερώτημα. Το κομβικό σημείο στην μετα-ανάλυση έθεσε ο *Eugene Glass* το 1976, ο οποίος κατάφερε να αποδώσει ένα τελικό συμπέρασμα από την σύνθεση συνόλου ανεξάρτητων δεδομένων. Για την εφαρμογή της αρχικά οφείλουμε να καθορίσουμε το αντικείμενο της μελέτης και το ερώτημα προς απάντηση και έπειτα από αναζήτηση στην βιβλιογραφία να επιλεγεί η καταλληλότερη μέθοδος προσέγγισης (Normand, 1999). Στην περίπτωση μας εξάγουμε τα πρωταρχικά δεδομένα από τις μελέτες που θα επιλεγθούν και μέσω των μεθόδων που περιεγράφηκαν στοχεύουμε στην ομοιομορφία αυτών ώστε να προχωρήσουμε στον καθορισμό του μεγέθους επίδρασης. Ως μέγεθος επίδρασης χρησιμοποιείται η διαφορά μέσων τιμών, ο συντελεστής συσχέτισης και το odds ratio που εκφράζει την πιθανότητα ένα γεγονός να συμβεί προς την πιθανότητα να μην συμβεί (Bagkos, 2015).

Η δύναμη της μετα-ανάλυσης εξαρτάται από τον αριθμό των μελετών και την μέθοδο που χρησιμοποιείται για τον συνδυασμό των μεμονωμένων εκτιμήσεων μεγέθους επίδρασης. Για τον συνδυασμό των εκτιμήσεων αυτών υπάρχουν δύο μοντέλα, των σταθερών και των τυχαίων επιδράσεων (*fixed* και *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. Στην περίπτωση αυτή κάθε μελέτη έχει διαφορετικό μέγεθος επίδρασης και διακύμανση ενώ κάθε δείγμα έχει μέγεθος επίδρασης κατανεμημένο σύμφωνα με την ανάλογη μέση τιμή και διακύμανση.

Για την ανάλυση δεδομένων από μικροσυστοιχίες χρησιμοποιούνται δύο κύριες κατηγορίες μεθόδων: αυτές που αφορούν στατιστική ανάλυση δεδομένων και αυτές που σχετίζονται με την ομαδοποίηση των δεδομένων (Ramamy et al., 2008). Βάση στατιστικής ανάλυσης χρησιμοποιήθηκε το μοντέλο τυχαίων επιδράσεων (Choi et al., 2003), που χρησιμοποιεί την τυποποιημένη διαφορά μέσων τιμών για δύο καταστάσεις και τον παράγοντα *t-squared* για να αναδείξει τα στατιστικώς σημαντικά διαφορικά εκφρασμένα γονίδια. Κατ' αυτόν τον τρόπο προκύπτει μια τιμή *p-value* για κάθε γονίδιο. Για την απαλοιφή ψευδο-θετικών (*false positive*) σημαντικών γονιδίων εκτελούνται πολλαπλές μέθοδοι διόρθωσης (Moreau et al., 2003).

Η εφαρμογή της μεθόδου μετα-ανάλυσης έγινε στο στατιστικό πρόγραμμα STATA (StataCorp, 2013) όπου υλοποιήθηκαν και οι ανάλογες μέθοδοι διόρθωσης για πολλαπλές συγκρίσεις:

- **Bonferroni**: το p-value κάθε γονιδίου πολλαπλασιάζεται με τον συνολικό αριθμό γονιδίων της μικροσυστοιχίας (Dudoit and Speed, 2000). Αν το διορθωμένο p-value βρίσκεται κάτω από το όριο σφάλματος, το γονίδιο χαρακτηρίζεται ως στατιστικά σημαντικό.
- **Bonferroni Step-Down (Holm)**: για τα p-value κάθε γονιδίου γίνεται ταξινόμηση από το μικρότερο στο μεγαλύτερο (Holm, 1979) και αποκτά δείκτη i . Έπειτα υπολογίζεται το διορθωμένο p-value από τον τύπο:

$$p_{corr(i)} = (n - i) * p_i$$

όπου n ο συνολικός αριθμός γονιδίων της μικροσυστοιχίας.

- **Benjamini-Hochberg False Discovery Rate (FDR)**: αποτελεί λιγότερο ισχυρή μέθοδο διόρθωσης αλλά χρησιμοποιείται σε δεδομένα μικροσυστοιχιών καθώς αποδίδει μια καλή ισορροπία μεταξύ στατιστικά σημαντικών και ψευδο-θετικών γονιδίων (Benjamini and Hochberg, 1995). Αρχικά γίνεται ταξινόμηση των p-values από το μικρότερο στο μεγαλύτερο, το μεγαλύτερο διατηρεί την τιμή του ενώ τα υπόλοιπα υπολογίζονται ως:

$$p_{corr(i)} = n / (n - i) * p_i.$$

- **Holland**: Ομοίως με την μέθοδο Holm, έπειτα από ταξινόμηση, το μικρότερο p-value πολλαπλασιάζεται με $(n - 1)$ ενώ το υπόλοιπα προκύπτουν ως:

$$p_{corr(i)} = (n - i + 1) * p_i$$

(Holland and Copenhaver, 1987)

- **Sidak**: κατά την οποία το διορθωμένο p-value υπολογίζεται από τον τύπο:

$$p_{corr(i)} = 1 - (1 - p_i)^{1/n}$$

(Šidák, 1967)

Ως μέγεθος επίδρασης χρησιμοποιήθηκε η διαφορά των μέσων τιμών μεταξύ των δειγμάτων ελέγχου και αναφοράς η οποία μπορεί να είναι τυποποιημένη (*standardized*) (Εξίσωση 9) ή μη τυποποιημένη (*unstandardized*) (Εξίσωση 8) (Thakkinstian et al., 2005). Στην πρώτη περίπτωση, ο παρονομαστής στον υπολογισμό της τυποποιημένης διαφοράς μέσων τιμών αποτελεί την αντίστοιχη ενοποιημένη τυπική απόκλιση (*standard deviation*) από τις δύο καταστάσεις (Εξίσωση 10). Για υπολογισμό αυτής οφείλουν να υπολογιστούν και οι τυπικές αποκλίσεις για κάθε κατάσταση (Εξίσωση 11). Έπειτα απαιτείται ο υπολογισμός της απόκλισης της τυποποιημένης μέσης τιμής της διαφοράς (Εξίσωση 12).

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

Εξίσωση 8: Μη-τυποποιημένη διαφορά μέσω των τιμών (unstandardized mean difference).

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

Εξίσωση 9: Τυποποιημένη διαφορά μέσω των τιμών (standardized mean difference).

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

Εξίσωση 10: Ενοποιημένη τυπική απόκλιση για δύο καταστάσεις. Τα n υποδηλώνουν τον αριθμό των δειγμάτων για τις δύο καταστάσεις.

$$sd_{ji} = \sqrt{\frac{\sum |x_{ji} - \bar{x}_{ji}|^2}{n_{ji} - 1}}$$

Εξίσωση 11: Τυπική απόκλιση για κάθε κατάσταση μεμονωμένα, όπου ο παρονομαστής εκφράζει τον βαθμό ελευθερίας της μελέτης καθώς το n υποδηλώνει τον αριθμό των δειγμάτων.

$$V_d = \frac{n_1 + n_2}{n_1 * n_2} + \frac{d_i^2}{2(n_1 + n_2)}$$

Εξίσωση 12: Απόκλιση της τυποποιημένης μέσης τιμής της διαφοράς (standardized mean difference deviation) για μια μελέτη.

Η τυποποιημένη διαφορά τείνει να υπερεκτιμήσει την απόλυτη τιμή της τελικής διαφοράς έκφρασης (diff) που επιθυμούμε να υπολογίσουμε, όταν έχουμε λίγα δείγματα. Για την εξάλειψη αυτού του φαινομένου απαιτείται μία διόρθωση που αποδίδει μια πιο καθαρή εκτίμηση για τιμή της διαφορικής έκφρασης, η οποία είναι γνωστή ως *g του Hedges*. Για την μετατροπή (Εξίσωση 14) από την τυποποιημένη διαφορά d στο g χρησιμοποιείται ένας διορθωτικός παράγοντας (Εξίσωση 13). Χρησιμοποιώντας το μοντέλο των σταθερών επιδράσεων υπολογίζουμε έναν συντελεστή που καθορίζει την βαρύτητα (*weight*) της σύγκρισης σε κάθε μελέτη (Εξίσωση 16) χρησιμοποιώντας το g ως μέγεθος επίδρασης και V_g την απόκλιση αυτού για κάθε μελέτη (Εξίσωση 15).

$$J = 1 - \frac{3}{4df - 1}$$

Εξίσωση 13: Διορθωτικός παράγοντας J για μετατροπή στο g του Hedges, όπου df ο βαθμός ελευθερίας (n1+n2-2) για ένα γονίδιο.

$$g_{study} = d * J$$

Εξίσωση 14: Μέγεθος επίδρασης κάθε μελέτης (Hedge's g).

$$V_g = J^2 * V_d$$

Εξίσωση 15: Απόκλιση του μεγέθους επίδρασης.

$$W_{study} = \frac{1}{V_g}$$

Εξίσωση 16: Το weight κάθε μελέτης έπειτα από μετατροπή στο g του Hedges και χρήση αυτού ως μέγεθος επίδρασης.

Μέχρι στιγμής η διαδικασία υπολογισμού της διαφοράς στην έκφραση ενός γονιδίου μεταξύ δύο καταστάσεων γίνεται με το μοντέλο των σταθερών επιδράσεων. Υπολογίζοντας τον παράγοντα *t-squared* (Εξίσωση 17) και ενσωματώνοντάς αυτόν στο weight κάθε μελέτης, θα μεταβούμε στο μοντέλο των τυχαίων επιδράσεων.

$$T^2 = \frac{Q - df}{C}$$

Όπου:

$$Q = \left\{ \sum W_{study} g_{study}^2 \right\} - \frac{(\sum W_{study} g_{study})^2}{\sum W_{study}}$$

$$C = \left\{ \sum W_{study} \right\} - \frac{\sum W_{study}^2}{\sum W_{study}}$$

Εξίσωση 17: Υπολογίζεται ως *t-squared* η διακύμανση των μεγεθών επίδρασης των αποτελεσμάτων στον πληθυσμό των μελετών.

Έτσι υπολογίζονται οι τροποποιημένες τιμές για την απόκλιση του μεγέθους επίδρασης και το *weight* κάθε μελέτης (Εξίσωση 18, Εξίσωση 19) για το μοντέλο των τυχαίων επιδράσεων.

$$V_{tot,study} = T^2 + V_g$$

Εξίσωση 18: Νέα απόκλιση του μεγέθους επίδρασης σε συνδυασμό με το *t-squared*.

$$W_{study}^* = \frac{1}{V_{tot,study}}$$

Εξίσωση 19: Το νέο weight κάθε μελέτης με την επίδραση του *t-squared*.

Έτσι προκύπτει η τελική τιμή διαφοράς έκφρασης για την σύγκριση δύο καταστάσεων για ένα γονίδιο από όλες τις μελέτες στις οποίες εμφανίστηκε (Εξίσωση 20). Επίσης από τον υπολογισμό της απόκλισης (Εξίσωση 21) για την τιμή αυτή προκύπτει το τυπικό σφάλμα (Εξίσωση 22).

$$diff = \frac{\sum W_{study}^* V_g}{\sum W_{study}^*}$$

Εξίσωση 20: Τελική τιμή διαφοράς έκφρασης για την σύγκριση δύο καταστάσεων για ένα γονίδιο από όλες τις μελέτες στις οποίες εμφανίστηκε.

$$V_{diff} = \frac{1}{\sum W_{study}^*}$$

Εξίσωση 21: Απόκλιση της τελικής τιμής διαφοράς έκφρασης για ένα γονίδιο σε μία σύγκριση.

$$SE = \sqrt{V_{diff}}$$

Εξίσωση 22: Τυπικό σφάλμα (standard error) για την τιμή της διαφοράς έκφρασης, για την σύγκριση ενός γονιδίου, για όλες τις μελέτες όπου εμφανίστηκε.

Από τον υπολογισμό του τυπικού σφάλματος λαμβάνεται η τιμή για το *Z-score* (Εξίσωση 23) για τον έλεγχο της μηδενικής υπόθεσης, που θα καθορίσει την τιμή *p_value* (Εξίσωση 24) για την αξιοπιστία της τιμής της διαφορικής έκφρασης που υπολογίσαμε (Borenstein et al., 2009).

$$Z_{score} = \frac{diff}{SE}$$

Εξίσωση 23: Υπολογισμός της τιμής *Z* για την μηδενική υπόθεση.

$$p_{value} = [1 - \Phi(\pm abs(Z_{score}))] * 2$$

Εξίσωση 24: Υπολογισμός *p-value*, όπου $\Phi(Z)$ η τυπική κανονική αθροιστική κατανομή, η οποία εφαρμόζεται στο Excel ως **NORMSDIST(Z)** θεωρώντας μέσο (mean) το μηδέν και τυπική απόκλιση 1.

Για να διεξαχθούν τα βέλτιστα αποτελέσματα και να αποφευχθεί η πιθανότητα υπερεκτίμησης ή υποεκτίμησης αυτών, απαραίτητη προϋπόθεση είναι η εκτενέστερη αναζήτηση στη βιβλιογραφία για την εύρεση όλων των μελετών που ταιριάζουν στο προφίλ της δικής μας ανάλυσης. Τέλος έμφαση πρέπει να δοθεί στην εξέταση του βαθμού ετερογένειας μεταξύ των μεμονωμένων μελετών καθώς η ανομοιογένεια των δειγμάτων μπορεί να οδηγήσει επίσης σε υπερεκτίμηση ή υποεκτίμηση των αποτελεσμάτων που θα διεξαχθούν (Thakkinstian et al., 2005). Στοιχεύουμε λοιπόν με την βοήθεια της μεθόδου της μετα-ανάλυσης να εκμεταλλευτούμε τον μεγάλο όγκο πειραματικών δεδομένων ώστε να ξεχωρίσουμε στατιστικώς σημαντικά, διαφορικώς εκφραζόμενα γονίδια τα οποία δεν θα εμφανίζονταν στην περίπτωση μίας απλούστερης άμεσης επεξεργασίας.

5 ΜΕΘΟΔΟΣ ΚΑΙ ΔΙΑΔΙΚΑΣΙΑ

Για την παρούσα εργασία χρησιμοποιήθηκε ο υπολογιστής HP15 -AC121nv (Intel Core i7 6500U/4 GB/CPU 2.50GHz) της *Hewlett Packard*, με περιβάλλον Windows 10. Σε αυτόν εγκαταστάθηκαν και χρησιμοποιήθηκαν: η γλώσσα προγραμματισμού της R (Brazma, 2009; Durinck et al., 2005; Gentleman et al., 2004; Oliver, 2003), το RStudio (RStudio, 2012) καθώς και το πακέτο της *limma* (Smyth, 2005) από το Bioconductor (Gentleman et al., 2004). Η βάση δεδομένων χτίστηκε σε SQL ενώ κάποια από τα scripts που χρειάστηκε να δημιουργηθούν έγιναν σε γλώσσα προγραμματισμού PHP (PHP Documentation Group, 1997). Η μετα-ανάλυση διεξήχθη με το στατιστικό πρόγραμμα STATA (StataCorp, 2013), ενώ για την περαιτέρω επεξεργασία των αποτελεσμάτων χρησιμοποιήθηκαν: το ολοκληρωμένο σύστημα εξόρυξης δεδομένων βασισμένο στον διαδικτυακό χώρο WebGestalt (Wang et al., 2013; Zhang et al., 2005) και η βάση δεδομένων STRING v 10.0 (Szklarczyk et al., 2015).

5.1 Αναζήτηση Πρωτογενών Δεδομένων

Στην μελέτη αυτή επιλέχθηκαν πειράματα μικροσυστοιχιών που χρησιμοποιούν δείγματα από ανθρώπινο βιολογικό υλικό, τα οποία έχουν ακτινοβοληθεί με χαμηλές δόσεις και περιλάμβαναν μη ακτινοβολημένα δείγματα (*control samples*) προκειμένου να είναι δυνατή η σύγκριση και άρα η εξαγωγή διαφορικά εκφραζόμενων γονιδίων. Αν και το όριο για τις χαμηλές δόσεις είναι $<0.1\text{Gy}$, οι μελέτες που χρησιμοποιήσαμε όριζαν ως όριο τα 0.5Gy και για το λόγο αυτό ενοποιήσαμε τα δείγματα που εκτέθηκαν σε δόση ιοντίζουσας ακτινοβολίας έως 0.5Gy θεωρώντας την ως έκθεση σε χαμηλή δόση. Εφόσον μια μελέτη περιείχε επιπλέον δείγματα που έχουν εκτεθεί σε υψηλές δόσεις ακτινοβολίας ($>0.5\text{Gy}$) θα συμπεριλήφθηκαν στην ανάλυση αυτή. Απορρίφθηκαν πειράματα που δεν περιείχαν δείγματα ελέγχου καθώς δεν θα υπήρχε δυνατότητα σύγκρισης της γονιδιακής έκφρασης με τα φυσιολογικά επίπεδα και άρα δεν θα ήταν δυνατός ο υπολογισμός της διαφορικής έκφρασης. Επίσης αγνοήθηκαν δείγματα που είχαν υποστεί χημική επεξεργασία σε συνδυασμό με ιοντίζουσες ακτινοβολίες ή έκθεση σε ακτινοβολία κατά διαστήματα (*chronic radiation*) ώστε να επιτευχθεί η μέγιστη δυνατή ομοιομορφία των δεδομένων προς ανάλυση. Τέλος δεν συμπεριλήφθηκαν στην ανάλυση, δείγματα που λήφθηκαν κάτω από 4 ώρες μετά την έκθεση, καθώς σε αυτό το διάστημα οι επιδράσεις της ακτινοβολίας δεν είναι ενδεικτικές για τα τελικά αποτελέσματα στον ιστό και την σκιαγράφηση των βιολογικών μηχανισμών απόκρισης.

5.1.1 Gene Expression Omnibus (GEO)

Η αναζήτηση για τα πρωταρχικά μας δεδομένα έγινε στο *Gene Expression Omnibus*, που αποτελεί δημόσιο αποθετήριο δεδομένων γονιδιακής έκφρασης (Barrett et al., 2011; Clough and Barrett, 2016), συμμορφωμένων προς τις απαιτήσεις του MIAME (*Minimal Information About Microarray Experiments*) (Brazma, 2009; Oliver, 2003). Στο GEO τα δεδομένα αρχειοθετούνται ως εξής: κάθε πειραματική πλατφόρμα έχει ένα μοναδικό αριθμό GPL. Κάθε δείγμα έχει ένα μοναδικό αριθμό GSM και ανήκει σε μία πλατφόρμα (GPL). Κάθε σειρά δειγμάτων έχει ένα μοναδικό αριθμό GSE και αποτελεί σύνολο από ένα ή περισσότερα δείγματα (GSMs) που ανήκουν σε μία ή περισσότερες πλατφόρμες (GPLs). Το ίδιο GSM μπορεί να ανήκει σε περισσότερες από μία σειρές (GSEs). Τέλος κάθε σύνολο δεδομένων έχει ένα μοναδικό αριθμό GDS και αποτελεί μια επιμελημένη συλλογή περισσότερων του ενός δειγμάτων (GSMs) που ανήκουν σε μία πλατφόρμα (GPL) και μία σειρά (GSE.)

Η αναζήτησή έγινε ως προς σειρές δειγμάτων. Ξεκινήσαμε λοιπόν με την αναζήτηση: ["radiation response"[All Fields] OR "response to radiation"[All Fields]) AND "Homo sapiens"[Organism]] ή οποία έδωσε 76 σειρές δεδομένων. Κάθε σειρά-μελέτη περιείχε βασικό σχολιασμό των δειγμάτων, συμπεριλαμβανομένων των πειραματικών παραγόντων και των τιμών τους (π.χ. ακτινοβολία, δόση), τον πειραματικό σχεδιασμό, τον τύπο της μικροσυστοιχίας και τα βασικά πρωτόκολλα εργαστηριακής επεξεργασίας. Ελέγχθηκαν όλες μία προς μία για την καταλληλότητά τους. Για να διαχωρίσουμε πλήρως τις κατάλληλες προς επεξεργασία μελέτες χρησιμοποιήσαμε επιπλέον στην αναζήτηση τους *MeSH Terms* (Dhammi and Kumar, 2014), που αποτελούν «έξυπνους» όρους στην μηχανή αναζήτησης του GEO. Επειδή με την χρήση του ["low dose radiation"] δεν βρέθηκε καμία εισαγωγή στο αποθετήριο καταλήξαμε στην αναζήτηση: ["whole-body irradiation"[MeSH Terms] AND "Homo sapiens"[Organism]] που έδωσε 9 σειρές. Πάλι ελέγχθηκαν όλες οι σειρές δεδομένων, σύμφωνα με τη διαδικασία που ακολουθήθηκε και τον τύπο των δεδομένων, ώστε να ταιριάζουν στο προφίλ της μελέτης αυτής (Πίνακας 3).

Πίνακας 3: Αναλυτικά στοιχεία για τις μελέτες που επιλέχθηκαν και αναλύθηκαν με τις αντίστοιχες πλατφόρμες που τους αντιστοιχούν.

GSE#	Low Dose [Gy]	High Dose [Gy]	Sample	IR Type
GSE8917	0.5	2, 5, 8	Αίμα	Cs-137 (gamma-rays)
GSE12435	0.5	-	Ινοβλάστες Πνεύμονα IMR-90	alpha-particles
GSE16935	0.1	2.5	Κερατινοκύτταρα EPI-200	Low LET protons
GSE20629	0.1	2.5	Κερατινοκύτταρα EPI-200	Low LET protons
GSE23901	0.5	-	Κερατινοκύτταρα EPI-200	alpha-particles
GSE29344	0.1	1	EpiDermFT	X-rays
GSE52918	0.1	2	Ινοβλάστες προστάτη	Cs-137 (gamma-rays)
GSE59861	0.05	2	Ινοβλάστες δέρματος AG1522	X-rays
GSE6978	0.05, 0.50	-	Λεμφοκύτταρα	Co-60 (gamma-rays)
GSE#	Platform			
GSE8917	Agilent-012391 Whole Human Genome Oligo Microarray G4112A (Feature Number version)		GPL1708	
GSE12435	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)		GPL6480	
GSE16935	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)		GPL6480	
GSE20629	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)		GPL6480	
GSE23901	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Feature Number version)		GPL6480	
GSE29344	Illumina HumanRef-8 v3.0 expression beadchip		GPL6883	
GSE52918	Illumina HumanHT-12 V4.0 expression beadchip		GPL10558	
GSE59861	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)		GPL16043	
GSE6978	CEA - DSV/DRR/SGF : Human 25k oligo v1.0		GPL4803	

5.1.2 Πειραματικός Σχεδιασμός των Επιλεγμένων Σειρών

- **GSE8917:** μελέτη της γονιδιακής έκφρασης σε ανθρώπινα κύτταρα αίματος μετά από έκθεση σε χαμηλές και υψηλές δόσεις ακτινοβολίας- γ . Διεξήχθησαν πέντε διαφορετικά πειράματα από ξεχωριστούς δότες ενώ τα δείγματα λήφθηκαν στις 6 και 24 ώρες μετά την έκθεση. Τέλος δείγματα λήφθηκαν και από μη ακτινοβολημένα κύτταρα (Paul and Amundson, 2008).
- **GSE12435:** τα δείγματα που χρησιμοποιήσαμε από αυτή την μελέτη δημιουργήθηκαν από κυτταρική σειρά που προέρχεται από ινοβλάστες πνεύμονα ανθρώπινου εμβρύου (IMR-90), ακτινοβολημένα με σωματίδια- α χαμηλής δόσης (0.5Gy) και τα αντίστοιχα μη ακτινοβολημένα (Ghandhi et al., 2008).
- **GSE16935:** τα δείγματα που χρησιμοποιήσαμε από αυτή την μελέτη λήφθηκαν για 6 και 24 ώρες έπειτα από έκθεση σε χαμηλές (0.1Gy) και υψηλές (2.5Gy) δόσεις ακτινοβολίας low LET πρωτονίων και τα αντίστοιχα μη ακτινοβολημένα. Έγιναν τρεις επαναλήψεις για κάθε κατάσταση. Χρησιμοποιήθηκε το EPI-200, ένα μοντέλο τρισδιάστατου ιστού, που μιμείται τη δομή και τη λειτουργία της ανθρώπινης επιδερμίδας και συνίσταται από κερατινοκύτταρα (Mezentsev and Amundson, 2011).
- **GSE20629:** στην μελέτη αυτή χρησιμοποιήθηκαν τα ίδια πειραματικά δεδομένα με τα παραπάνω και τα αποτελέσματά της παρατέθηκαν στην ίδια δημοσίευση (Mezentsev and Amundson, 2011).
- **GSE23901:** το πείραμα διεξήχθη με σκοπό την μελέτη της γονιδιακής έκφρασης έπειτα από ακτινοβολήση, σε τρισδιάστατο μοντέλο ιστού δέρματος (EPI-200). Τα δείγματα που χρησιμοποιήσαμε από αυτή την μελέτη ήταν όσα λήφθηκαν για 16 ώρες έπειτα από έκθεση σε χαμηλές δόσεις (0.5Gy) ακτινοβολίας πρωτονίων (ακτινοβολία- β), low LET και τα αντίστοιχα μη ακτινοβολημένα. Έγιναν τρεις επαναλήψεις για κάθε κατάσταση.
- **GSE29344:** το μοντέλο που χρησιμοποιήθηκε στο πείραμα αυτό, είναι ανακατασκευασμένος δερματικός ιστός, ο οποίος αποτελείται από κερατινοκύτταρα που αποτελούν την επιδερμική στιβάδα και ινοβλάστες που αποτελούν το δερματικό στρώμα του δέρματος (EpiDermFT). Οι ιστοί ακτινοβολήθηκαν με ακτίνες-X 0.10 και 100 cGy. Χρησιμοποιήθηκαν τα δείγματα που συλλέχθηκαν στις 8 και 24 ώρες μετά την ακτινοβολήση. (Ray et al., 2012).

- **GSE52918:** δείγματα από καλλιέργειες ινοβλαστών ανθρώπινου προστάτη συλλέχθηκαν σε 24 ώρες μετά από έκθεση σε 10cGy ή 200cGy ακτίνων-γ. Τα δείγματα ελέγχου συλλέχθηκαν ταυτόχρονα. Χρησιμοποιήθηκαν συνολικά 4 ανεξάρτητα πειραματικά αντίγραφα για κάθε δείγμα (McDonald et al., 2014).
- **GSE59861:** στη μελέτη αυτή χρησιμοποιήθηκαν δείγματα από την κυτταρική σειρά AG1522 που απαρτίζεται από ινοβλάστες ανθρώπινου δέρματος. Επιλέχθηκαν για την ανάλυση αυτή όσα εκτέθηκαν σε 5cGy και 2Gy ακτίνων-Χ και συλλέχθηκαν 6, 12 και 24 ώρες μετά την έκθεση και τα αντίστοιχα μη ακτινοβολημένα (Hou et al., 2015).
- **GSE6978:** στο πείραμα αυτό μελετήθηκε η διαφορική έκφραση γονιδίων σε λεμφοκύτταρα έπειτα από έκθεση σε χαμηλές δόσεις (0.05-0.5Gy) ακτινοβολίας-γ. Χρησιμοποιήθηκαν μόνο τα δείγματα που λήφθηκαν 24 ώρες μετά την ακτινοβολήση και η μελέτη έγινε με διχρωματική μικροσυστοιχία (Gruel et al., 2008).

Λόγω της ύπαρξης διαφορετικών πλατφορμών στις μελέτες αυτές, η αρχική επεξεργασία διόρθωσης υποβάθρου και κανονικοποίησης έως την επίτευξη ομοιογένειας των δεδομένων για την εισαγωγή αυτών στην βάση δεδομένων, διαχωρίστηκε σε δύο μέρη. Στην εργασία αυτή επεξεργάστηκαν: οι πλατφόρμες της Agilent και η CEA - DSV/DRR/SGF: Human 25k oligo v1.0. Έπειτα από την εισαγωγή του συνόλου των επεξεργασμένων δεδομένων από όλες τις σειρές δεδομένων στην βάση «microarray» εξήγαμε τις σταθμισμένες εκφράσεις και έγινε διαχωρισμός των δειγμάτων με βάση και το χρονικό πλαίσιο που αυτά συλλέχθηκαν έπειτα από την ακτινοβολήση (Script 6).

Σύμφωνα με την κατανομή των δειγμάτων από τις μελέτες που χρησιμοποιήθηκαν και την γενικότερη βιβλιογραφία, επιλέξαμε δύο διαστήματα: μεταξύ τεσσάρων και δεκαέξι ωρών και από δεκαέξι έως είκοσι τέσσερις ώρες. Σημειώνεται ότι όταν επέλθει το διάστημα των 24 ωρών, το κύτταρο έχει επιστρέψει σχεδόν πλήρως στην φυσιολογική του κατάσταση μέσω μηχανισμών επιδιόρθωσης και για αυτό δεν θα παρατηρούνταν σημαντική διαφορική έκφραση σε σχέση με τα δείγματα ελέγχου. Η επιλογή των ωρών έγινε ώστε η ποσότητα των δειγμάτων να είναι ικανοποιητική και για τις δύο ομάδες, για την βέλτιστη εξαγωγή αποτελεσμάτων στατιστικά αλλά και βιολογικά. Στην εργασία αυτή αναλύονται τα δείγματα για το χρονικό πλαίσιο 4-16 ωρών (Πίνακας 4), που αποτελεί ένα πρωταρχικό στάδιο σκιαγράφησης και πρόβλεψης των επιδράσεων της ιοντίζουσας ακτινοβολίας.

Με εφαρμογή Μπεϋζιανής ανάλυσης με χρήση της R για τον υπολογισμό των p-values έγινε ένας έλεγχος για τα γονίδια που προκύπτουν ως στατιστικά σημαντικά διαφορεκώς εκφραζόμενα. Στη συνέχεια προχωρήσαμε σε μετα-ανάλυση.

Πίνακας 4: Λίστα δειγμάτων που χρησιμοποιήθηκαν στην παρούσα εργασία με την αντίστοιχη κατάσταση (Control-Low-High) και την αντίστοιχη μελέτη στην οποία ανήκουν.

Study	Condition	Sample
GSE12435	Control	GSM312307
GSE12435	Control	GSM312308
GSE12435	Control	GSM312309
GSE12435	Control	GSM312310
GSE12435	Low	GSM312299
GSE12435	Low	GSM312300
GSE12435	Low	GSM312301
GSE12435	Low	GSM312302
GSE16935	Control	GSM424533
GSE16935	Control	GSM424534
GSE16935	Control	GSM424535
GSE16935	Control	GSM424542
GSE16935	Control	GSM424543
GSE16935	Control	GSM424544
GSE16935	High	GSM424539
GSE16935	High	GSM424540
GSE16935	High	GSM424541
GSE16935	Low	GSM424536
GSE16935	Low	GSM424537
GSE16935	Low	GSM424538
GSE23901	Control	GSM589482
GSE23901	Control	GSM589483
GSE23901	Control	GSM589484
GSE23901	Low	GSM589485
GSE23901	Low	GSM589486
GSE23901	Low	GSM589487
GSE29344	Control	GSM725372
GSE29344	Control	GSM725373
GSE29344	Control	GSM725376
GSE29344	Control	GSM725377
GSE29344	Control	GSM725378
GSE29344	Control	GSM725379
GSE29344	Control	GSM725380
GSE29344	Control	GSM725381
GSE29344	Control	GSM725388
GSE29344	Control	GSM725389
GSE29344	High	GSM725392
GSE29344	High	GSM725393
GSE29344	Low	GSM725384
GSE29344	Low	GSM725385
GSE59861	Control	GSM1448246
GSE59861	Control	GSM1448247
GSE59861	Control	GSM1448248
GSE59861	Control	GSM1448249
GSE59861	High	GSM1448253
GSE59861	High	GSM1448254
GSE59861	High	GSM1448255
GSE59861	High	GSM1448256

GSE59861	High	GSM1448257
GSE59861	High	GSM1448258
GSE59861	Low	GSM1448265
GSE59861	Low	GSM1448266
GSE59861	Low	GSM1448267
GSE59861	Low	GSM1448268
GSE59861	Low	GSM1448269
GSE59861	Low	GSM1448270
GSE8917	Control	GSM225984
GSE8917	Control	GSM225989
GSE8917	Control	GSM225994
GSE8917	Control	GSM225999
GSE8917	Control	GSM226004
GSE8917	Control	GSM226009
GSE8917	Control	GSM226014
GSE8917	Control	GSM226019
GSE8917	Control	GSM226024
GSE8917	Control	GSM226029
GSE8917	High	GSM225986
GSE8917	High	GSM225987
GSE8917	High	GSM225988
GSE8917	High	GSM225991
GSE8917	High	GSM225992
GSE8917	High	GSM225993
GSE8917	High	GSM225996
GSE8917	High	GSM225997
GSE8917	High	GSM225998
GSE8917	High	GSM226001
GSE8917	High	GSM226002
GSE8917	High	GSM226003
GSE8917	High	GSM226006
GSE8917	High	GSM226007
GSE8917	High	GSM226008
GSE8917	Low	GSM225985
GSE8917	Low	GSM225990
GSE8917	Low	GSM225995
GSE8917	Low	GSM226000
GSE8917	Low	GSM226005

5.2 Επεξεργασία με Χρήση RStudio

Οι εντολές για το κατέβασμα των πρωτογενών δεδομένων στον υπολογιστή είναι:

```
>source("http://bioconductor.org/biocLite.R")
>biocLite("GEOquery")
>library(GEOquery)
>getGEOSuppFiles("GSE8917")
```

Ενώ η εντολή για την εγκατάσταση του πακέτου `limma` είναι:

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

5.2.1 Agilent Platform

Χρησιμοποιήθηκε η `cmd` (γραμμή εντολών) για την δημιουργία αρχείου `targets.txt` που περιλαμβάνει σε λίστα τα ονόματα των `.txt` αρχείων που περιέχουν τα πρωταρχικά δεδομένα της μελέτης για κάθε δείγμα, το οποίο είναι οριοθετημένο με `tab` (*tab delimited*) ώστε να διαβάζεται από την R, σε στήλες μορφής:

SlideNo	FileNames
1	GSM225984.txt
2	GSM225985.txt
3	GSM225986.txt

Καλώντας το πακέτο `limma` και ορίζοντας έναν φάκελο εργασίας (*directory*) έγινε ανάγνωση, μέσω του `targets.txt`, των πρωτογενών δεδομένων όπως τα κατεβάσαμε από το GEO. Σε αυτό το σημείο είναι βασικό να σημειωθεί πως όλες οι μελέτες μας που χρησιμοποιούν την πλατφόρμα της Agilent είναι μονοκάναλες (*single channel* ή *one color arrays*) και με το όρισμα `green.only=TRUE` στην τελευταία εντολή διευκρινίζεται ότι πρόκειται για μονοχρωματική μικροσυστοιχία.

GSE8917:

```
>library(limma)
>setwd("C:/Microarrays/Data_Normalized/8917[Agilent]")
>targets = readTargets("targets.txt")
>targetsfinal = targets[,c(2)]
>raw=read.maimages(targets,source="agilent",green.only=TRUE)
```

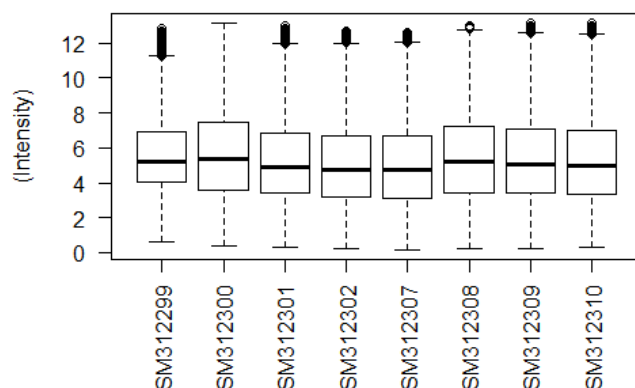
Σημειώνεται ότι κατά την ανάγνωση δεδομένων της πλατφόρμας Agilent που καθορίζεται στην εντολή `read.maimages()` με το όρισμα `source="agilent"`, τυπώνονται δύο στήλες εκ των οποίων η μία αποτελείται από τις τιμές γονιδιακής έκφρασης που λαμβάνονται από την διάμεσο (*Median Signal*), ενώ η άλλη αποτελείται από τις τιμές του σήματος υποβάθρου επίσης από την τιμή της διαμέσου (*BG Median*

Signal) για κάθε spot σε κάθε δείγμα, όπως συνηθίζεται για το υπόβαθρο σε όλες τις πλατφόρμες. Οι εντολές για την διόρθωση υποβάθρου και την κανονικοποίηση είναι:

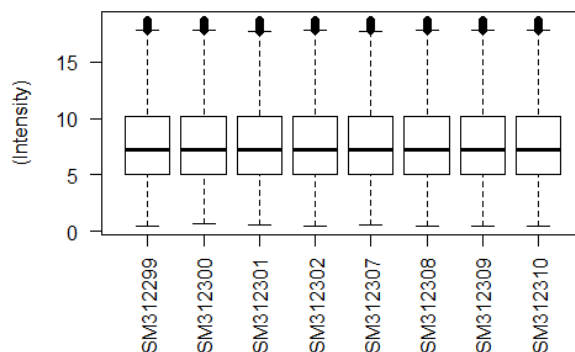
```
>BGcorrected=backgroundCorrect(raw,method="normexp",offset=1)
>Normalized=normalizeBetweenArrays(BGcorrected,method="quantile")
```

Οι εντολές για τον έλεγχο της κανονικοποίησης με χρήση του boxplot (Εικόνα 16, Εικόνα 17):

```
>boxplot(log(as.matrix(BGcorrected)),las=2,ylab="Intensity")
>boxplot(as.matrix(Normalized),las=2,ylab="Intensity")
```



Εικόνα 16: Boxplot για τις τιμές έκφρασης της μελέτης GSE8917 πριν την κανονικοποίηση.



Εικόνα 17: Boxplot μετά την κανονικοποίηση για την σειρά GSE8917.

```
># Δημιουργία τελικών αρχείων
>write.table(Normalized,file="norm_8917.txt",sep = "\t", na =
/"NA", dec = ".", quote = FALSE, col.names = TRUE, row.names = FALSE)
```

Ακολουθεί ο κώδικας για τις υπόλοιπες σειρές δεδομένων:

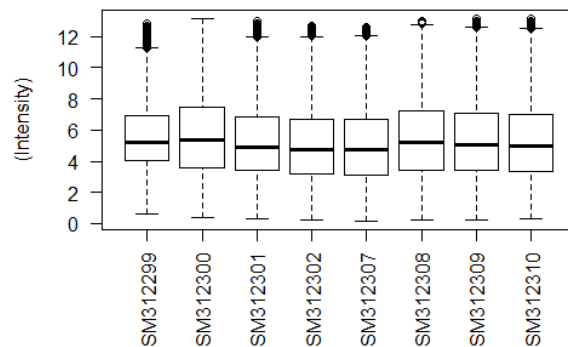
GSE12435:

```
>setwd("C:/Microarrays/Data_Normalized/12435[Agilent]")
>targets = readTargets("targets.txt")
```

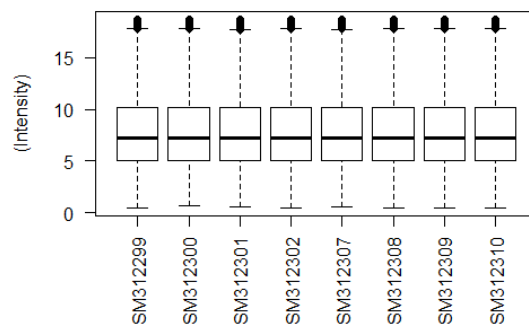
```

>targets = targets[,c(2)]
>raw=read.maimages(targets,source="agilent",green.only=TRUE)
>BGcorrected=backgroundCorrect(raw,method="normexp",offset=1)
>Normalized=normalizeBetweenArrays(BGcorrected,method="quantile")
>write.table(Normalized,file="norm_12435.txt", sep = "\t ", na =
/"NA", dec = ".", quote = FALSE, col.names = TRUE, row.names = FALSE)
># Έλεγχος κανονικοποίησης (Εικόνα 18, Εικόνα 19)
>boxplot(log(as.matrix(BGcorrected)),las=2,ylab="Intensity")
>boxplot(as.matrix(Normalized),las=2,ylab="Intensity")

```



Εικόνα 18: Βoxplot για τις τιμές έκφρασης της GSE12435 πριν την κανονικοποίηση.



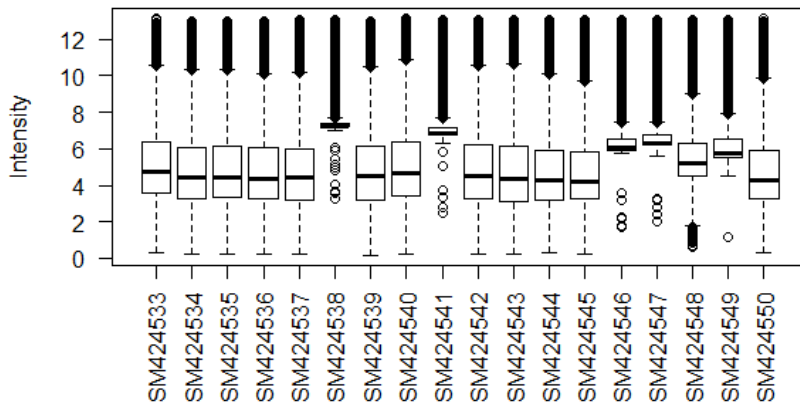
Εικόνα 19: Βoxplot μετά την κανονικοποίηση για την σειρά GSE12435.

GSE16935:

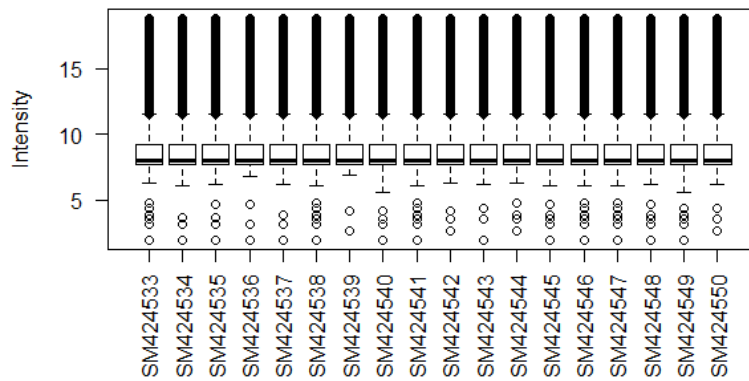
```

>setwd("C:/Microarrays/Data_Normalized/16935[Agilent]")
>targets = readTargets("targets.txt")
>targets = targets[,c(2)]
>raw=read.maimages(targets,source="agilent",green.only=TRUE)
>BGcorrected=backgroundCorrect(raw,method="normexp",offset=1)
>Normalized=normalizeBetweenArrays(BGcorrected,method="quantile")
>write.table(Normalized,file="norm_16935.txt", sep = "\t ", na =
/"NA", dec = ".", quote = FALSE, col.names = TRUE, row.names = FALSE)
># Έλεγχος με χρήση boxplot (Εικόνα 20, Εικόνα 21)
>boxplot(log(as.matrix(BGcorrected)),las=2,ylab="Intensity")
>boxplot(as.matrix(Normalized),las=2,ylab="Intensity")

```



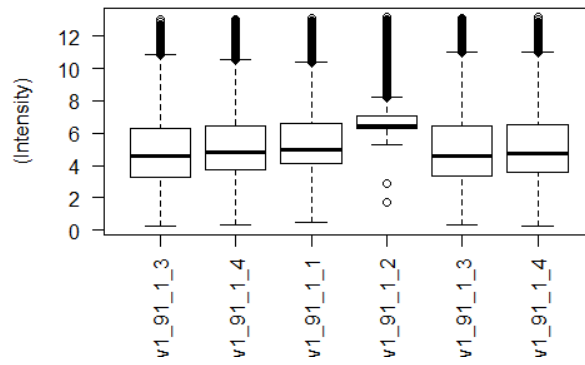
Εικόνα 20: Βοχplot για τις τιμές έκφρασης της GSE16935 πριν την κανονικοποίηση.



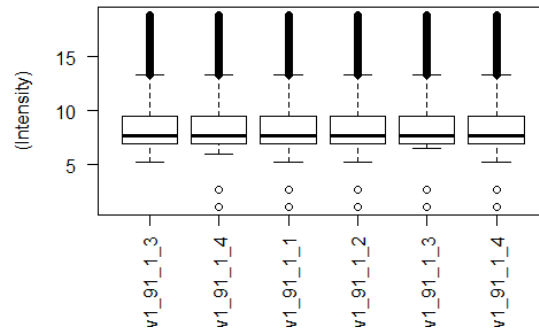
Εικόνα 21: Βοχplot μετά την κανονικοποίηση για την σειρά GSE16935.

GSE23901:

```
>setwd("C:/Microarrays/Data_Normalized/23901[Agilent\\]")
>targets = readTargets("targets.txt")
>targets = targets[,c(2)]
>raw=read.maimages(targets,source="agilent",green.only=TRUE)
>BGcorrected=backgroundCorrect(raw,method="normexp",offset=1)
>Normalized=normalizeBetweenArrays(BGcorrected,method="quantile")
>write.table(Normalized,file="norm_23901.txt",sep = "\t ", na =
/"NA", dec = ".", quote = FALSE, col.names = TRUE, row.names = FALSE)
># Έλεγχος με χρήση boxplot (Εικόνα 22, Εικόνα 23)
>boxplot(log(as.matrix(BGcorrected)),las=2,ylab="Intensity")
>boxplot(as.matrix(Normalized), las=2,ylab="Intensity")
```



Εικόνα 22: Βoxplot για τις τιμές έκφρασης της GSE23901 πριν την κανονικοποίηση.



Εικόνα 23: Βoxplot μετά την κανονικοποίηση για την σειρά GSE23901.

5.2.2 CEA - DSV/DRR/SGF: Human 25k oligo v1.0 (GenePix)

Η συγκεκριμένη πλατφόρμα είναι αυτοσχέδια και διχρωματική, παρόλα αυτά ο σαρωτής που χρησιμοποιήθηκε στις μικροσυστοιχίες του πειράματος είναι γνωστής τεχνολογίας GenePix, συνεπώς τα δεδομένα εξάγονται με γνωστή μορφή (*format*). Χρησιμοποιήθηκε αντιστοίχως το cmd για την δημιουργία αρχείου `targets.txt` που περιλαμβάνει σε λίστα τα ονόματα των `.gpr` αρχείων που περιέχουν τα πρωταρχικά δεδομένα της μελέτης για κάθε δείγμα, το οποίο έχει την εξής δομή:

Slide	FileNames
1	GSM160300.gpr
2	GSM160301.gpr

Ο αντίστοιχος κώδικας είναι:

```
>library(limma)
>setwd("C:/Microarrays/GSE6978[Geno]")
>targets = readTargets("targets.txt")
>targets = targets[,c(2)]
>RG = read.maimages(targets,source="genepix")
```

Κατά την ανάγνωση δεδομένων που έχουν συλλεχθεί από σαρωτή τύπου GenePix, όπως καθορίζεται στην εντολή `read.maimages()` με το όρισμα `source="genepix"`, τα δείγματα διαβάζονται δύο φορές για `$R(Red color/635nm/Cy5)` και `$G(Green color/532nm/Cy3)`. Εννοείται δηλαδή ότι `green.only=FALSE`. Τυπώνονται έτσι τέσσερις στήλες εκ των οποίων οι δύο αποτελούνται από τις τιμές γονιδιακής έκφρασης που λαμβάνονται από τον μέσο όρο (*Mean Signal*) για κάθε χρώμα αντίστοιχα, ενώ οι εναπομείναντες δύο αποτελούνται από τις τιμές του σήματος υποβάθρου από την τιμή της διαμέσου (*BG Median Signal*) για κάθε *spot* σε κάθε δείγμα, όπως γίνεται για όλες τις πλατφόρμες. Επιλέχθηκε ομοίως η μέθοδος `normexp` για την διόρθωση υποβάθρου (Ritchie et al., 2007) στην οποία όμως θέσαμε τιμή `offset=25` σύμφωνα με τον οδηγό του πακέτου `limma` (Smyth et al., 2002).

```
># Κανονικοποίηση που Εμπεριέχει Διόρθωση Υποβάθρου
>layout=list(ngrid.r=RG$printer$ngrid.r,
/ngrid.c=RG$printer$ngrid.c,nspt.r=RG$printer$nspt.r,
/nspt.c=RG$printer$nspt.c)
>MA=normalizeWithinArrays(RG,layout,method="printtiploess", span =
/0.4, iterations = 4, controlspots = NULL, bc.method="normexp",
/offset = 25)
>MA.f = normalizeBetweenArrays(MA,method="quantile")
```

Στην περίπτωση δικάναλικών μικροσυστοιχιών απαιτείται επιπλέον κανονικοποίηση μεταξύ των *spot* με την μέθοδο *loess* (Smyth and Altman, 2013) σε μία μικροσυστοιχία,

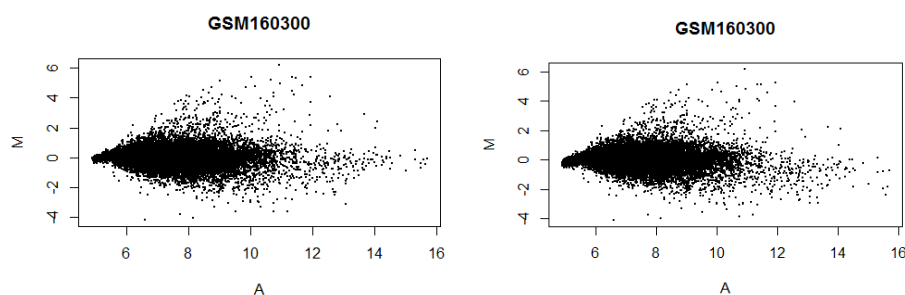
πριν προχωρήσουμε σε κανονικοποίηση ποσοστημορίου μεταξύ των μικροσυστοιχιών. Έπειτα από την κανονικοποίηση λαμβάνουμε τις τιμές M, δηλαδή την διαφορά στην έκφραση μεταξύ των δύο δειγμάτων-χρωμάτων και την συνολική φωτεινότητα A για κάθε spot (Εξίσωση 25).

$$M = \log_2(R) - \log_2(G) = \log_2\left(\frac{R}{G}\right)$$

$$A = \frac{\log_2(R) + \log_2(G)}{2} = \frac{\log_2(R * G)}{2}$$

Εξίσωση 25: Κανονικοποιημένες τιμές έκφρασης για διχρωματική μικροσυστοιχία.

Οι τιμές A δεν χρησιμεύουν στην ανάλυση μας. Για τιμές M=0 έχουμε ίση έκφραση, M=1 διπλάσια έκφραση του ενός χρώματος ως προς το άλλο κ.ο.κ. Για τις τιμές των ορισμάτων `offset` και `span` συμβουλευτήκαμε την βιβλιογραφία αλλά και πραγματοποιήσαμε ορισμένα *M*Plots (Εξίσωση 21) χρησιμοποιώντας την εντολή `plotMA()` για το ίδιο *slide* ώστε να καταλήξουμε στην βέλτιστη επιλογή για την τιμή του ορίσματος `span` στην εντολή της κανονικοποίησης (Jacoby, 2000; Smyth and Speed, 2003).

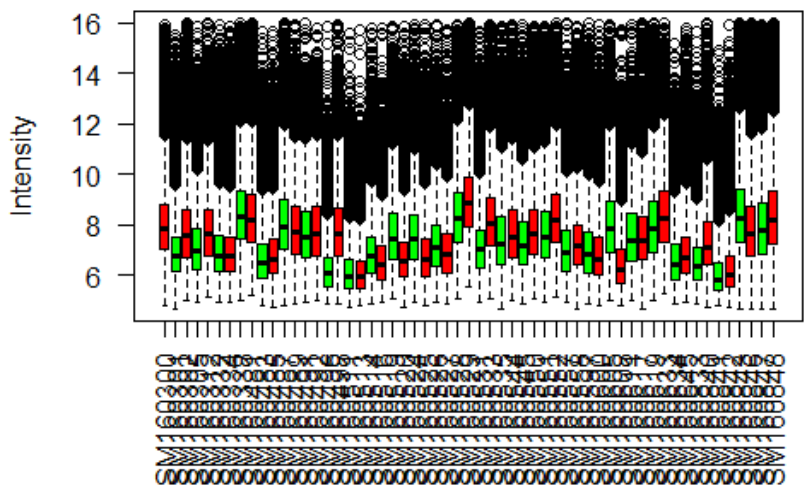


Εικόνα 24: α) *M*Plot με τιμή `span=0.4` και β) *M*Plot με τιμή `span=0.75` για το ίδιο *slide*.

```
>#Έλεγχος με χρήση boxplot για τα δεδομένα μετά την διόρθωση
υποβάθρου χωρίς κανονικοποίηση για τα 2 χρώματα
>RG_BG = backgroundCorrect(RG, method="normexp", offset=25,
\printer=RG$printer)
>boxplot(log2(data.frame(RG_BG$R, RG_BG$G, las=2,
\ylab="Intensity", col=rep(c("red", "green"))))
```

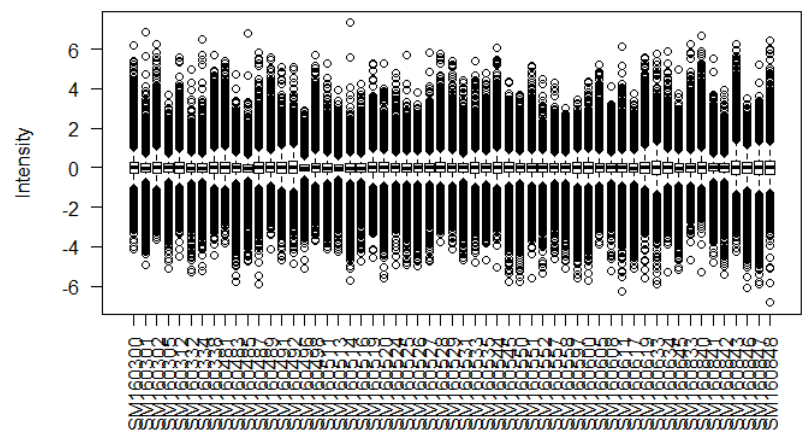
Καθώς τα δεδομένα από μικροσυστοιχίες δύο καναλιών μας δίνουν πληροφορίες για την διαφορική έκφραση μεταξύ των δύο χρωμάτων και συνεπώς εξάγουν σχετικές τιμές έκφρασης, σε αντίθεση με τις μονοχρωματικές που επιστρέφουν απόλυτες τιμές, θεωρήσαμε ότι δεν μπορούσε να επιτευχθεί ικανοποιητική ομοιογένεια και καθώς η μετα-ανάλυση απαιτεί απόλυτες τιμές έκφρασης αποφασίσαμε να αποκλείσουμε την

σειρά αυτή από την μετέπειτα επεξεργασία. Ενδεικτικά ελέγχθηκε η διαδικασία της κανονικοποίησης (Εικόνα 25, Εικόνα 26).



Εικόνα 25: Βoxplot για τις τιμές έκφρασης πριν την κανονικοποίηση που δίνουν για το GSE6978.

```
>#Δημιουργία boxplot για έλεγχο κανονικοποιημένων δεδομένων
>boxplot(data.frame(MA$M), las=2, ylab="Intensity")
```



Εικόνα 26: Βoxplot μετά την κανονικοποίηση για την σειρά GSE6978.

5.3 Περιγραφή της Βάσης Microarray

Η δημιουργία μιας βάσης δεδομένων είναι απαραίτητη όχι μόνο για την δυνατότητα ανάκλησης των επιθυμητών τιμών και δεδομένων μετά το πέρας της επεξεργασίας αλλά και για την βελτιστοποίηση της ανάλυσης και την διεξαγωγή μετα-ανάλυσης. Με την κατασκευή της βάσης εξαλείφεται αρχικά το ζήτημα των διαφορετικών πλατφορμών ως προς την αντιστοίχιση κάθε spot μέσω των μεταγράφων σε συγκεκριμένη ονοματολογία, ενώ προκύπτει μία σταθμισμένη τιμή έκφρασης στην περίπτωση που δύο spot αντιστοιχούν στο ίδιο γονίδιο αλλά έχουν διαφορετικό όνομα. Το διάγραμμα συσχέτισης των οντοτήτων της βάσης «microarray» παρατίθεται παρακάτω (Εικόνα 27). Ακολουθεί μια συνοπτική περιγραφή για το περιεχόμενο και τη χρησιμότητα κάθε πίνακα που συνθέτει την βάση δεδομένων:

- **ENSG**: αποτελείται από μία στήλη που περιέχει όλα τα ENSG IDs του ανθρώπινου οργανισμού όπως λήφθηκαν από το BioMart (Kasprzyk, 2011; Smedley et al., 2009) μέσω του προγράμματος περιήγησης γονιδιώματος Ensembl (Aken et al., 2017; Yates et al., 2016)
- **ENSG_HGNC**: αντιστοιχίζει όλα τα ENSG IDs του ανθρώπινου οργανισμού με τα αντίστοιχα ονόματα γονιδίων από την HUGO *Gene Nomenclature Committee* (Gray et al., 2015). Αποτελείται από δύο στήλες καθώς επεξεργάστηκε με χρήση `php` (Script 1) ώστε να διατηρηθούν ως πληροφορία μόνο τα ονόματα από το HGNC και τα ENSG IDs.
- **ENST**: αντιστοιχίζει τα μετάγραφα (ENS Transcripts) σε κάθε γονίδιο (ENSG IDs) ομοίως μέσω BioMart (Kasprzyk, 2011; Smedley et al., 2009).
- **ENST_Seq**: αντιστοιχεί τις αλληλουχίες από το BioMart στα μετάγραφα μέσω κατάλληλης διαχείρισης των αλληλουχιών. Οι αλληλουχίες εισάγονται σε μορφή FASTA (Pearson, 1994) και απαιτείται αφαίρεση του identifier (σύμβολο '>' στην 1η γραμμή του αρχείου) και επικόλληση των αλληλουχιών μεταξύ τους ανά μετάγραφο, το οποίο έγινε με χρήση `php` (Script 2).
- **nu**: περιέχει όλες τις αλληλουχίες των probes από chips Illumina και Agilent με τα αντίστοιχα μετάγραφα σε μορφή nuID (Du et al., 2007) που αναγνωρίζουν. Με το `script seq2id.php` (Script 3) μία αλληλουχία κωδικοποιείται και μετατρέπεται σε μοναδικό κωδικό, ο οποίος έχει μικρότερο μέγεθος ώστε να τρέχουν πιο γρήγορα οι εντολές της SQL και να είναι δυνατή η αντιστοίχιση από όλες τις πλατφόρμες.
- **ENST_nu**: περιέχει όλα τα transcripts (ENST IDs) και τις αντίστοιχες μοναδικές υπογραφές τους nuIDs (Du et al., 2007). Το SQL query για την δημιουργία

αυτού πίνακα αντιστοιχίζει τις ακολουθίες των *probes* από τον πίνακα nu με τις ακολουθίες των μεταγράφων (*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.

- **Probes:** περιέχει σε μία στήλη όλα τα probe IDs με βάση την πλατφόρμα στην οποία ανήκουν. Οι πληροφορίες εισήχθησαν από τα *manifest files* που περιέχουν τις απαραίτητες πληροφορίες κάθε μικροσυστοιχίας.
- **ProbesNU:** περιέχει όλα τα μοναδικά nuIDs από τα εμπορικά ονόματα των probes και τα αντίστοιχα τους probe IDs (Script 4). Συντίθεται μέσω επεξεργασίας όλων των αρχείων τύπου manifest των πλατφορμών και εξαγωγή των απαραίτητων στοιχείων, δηλαδή του ProbeID και της ακολουθίας του, η οποία μετατρέπεται παράλληλα σε μορφή nuID με το script seq2id.php
- **Expression:** περιέχει την κανονικοποιημένη τιμή έκφρασης που αντιστοιχεί σε κάθε featureID για κάθε GSM και για κάθε πλατφόρμα. Συμπληρώθηκε έπειτα από την επεξεργασία στο RStudio και με την βοήθεια κατάλληλων *parsers* γραμμένων σε R για κάθε πλατφόρμα.
- **Affymetrix:** Αντιστοιχίζει τα manifest files της πλατφόρμας της Affymetrix με τα αντίστοιχα ENSG IDs.
- **Features:** αντιστοιχίζει τα featureIDs κάθε πλατφόρμας με τα αντίστοιχα probeIDs τους. Γίνεται συντακτική ανάλυση (*parsing*) των αρχείων τύπου manifest και διατηρείται το featureID για τα Agilent ή το probeID για τα Illumina στην 1η στήλη, συμπληρώνεται το όνομα της πλατφόρμας στην 2η στήλη τέλος διατηρείται το probeID και εισάγεται στην 3η στήλη. Για τις πλατφόρμες Illumina και Affymetrix συνεπώς η 1η και η 3η στήλη είναι σε αντίθεση με την Agilent. Επίσης ειδικά στην περίπτωση της Agilent, η οντότητα αυτή είναι σημαντική διότι συναντώνται ίδιες αλληλουχίες σε διαφορετικά spot.
- **FeatureSet:** κάνει την αντιστοίχιση όλων των Feature IDs από κάθε platform με τα ENSG IDs τους
- **ProcExpr:** κάνει την αντιστοίχιση καθενός ENSG ID από κάθε GSM με ένα σταθμισμένο expression μέσω της μεθόδου *one-step Tukey* (Hoaglin et al., 1983; Huber and Ronchetti, 2009). Η μέθοδος αυτή αποτελεί μία διαδικασία πολλαπλών συγκρίσεων και στατιστικής δοκιμής για τη εύρεση μέσων τιμών που διαφέρουν σημαντικά μεταξύ τους. Συγκρίνει όλα τα πιθανά ζεύγη μέσων τιμών και βασίζεται σε τυποποιημένη κλίμακα κατανομής παρόμοια με την κατανομή του *t-test*.
- **Study:** περιέχει μια λίστα με τις μελέτες (GSEs).

- **Condition:** επιδεικνύει σε ποια κατάσταση ανήκει κάθε δείγμα (control – low - high).
- **Sample:** περιέχει τέσσερις στήλες με όλα τα GSM από κάθε πλατφόρμα και τις μελέτες (GSE) με τις αντίστοιχες καταστάσεις (*conditions*).
- **Platform:** αποτελείται από δύο στήλες με την πλατφόρμα και το αντίστοιχο GPL.
- **Average:** περιέχει για κάθε γονίδιο το GSE του και το μέγεθος δείγματος του GSE, τη μέση τιμή έκφρασης του, και την τυπική του απόκλιση σε κάθε κατάσταση. Το ερώτημα SQL (*SQL query*) εδώ δημιουργεί όλα τα περιγραφικά στατιστικά που χρειάζονται για την δημιουργία του πίνακα.
- **Bayes:** Όλα τα ENSG από κάθε GSE σε όλες τις καταστάσεις τους (σύγκριση σε ζεύγη καταστάσεων) και οι αντίστοιχες τιμές των p-values. Συμπληρώθηκε με χρήση του RStudio μέσω Μπεϋζιανής ανάλυσης.

Script 1: ENSG.edit.php για την διατήρηση των απαραίτητων πληροφοριών για τον πίνακα ENSG_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";
}
?>
```

Script 2: ENST_Seq.php για την διαχείριση των αλληλουχιών.

```
<?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";
?>
```

Script 3: seq2id.php. Κωδικοποίηση αλληλουχίας με σκοπό την γρηγορότερη λειτουργία των εντολών 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;
}
?>

```

Script 4: Probes.php. Επεξεργασία όλων των Manifest files των πλατφορμών και εξαγωγή των απαραίτητων στοιχείων (ProbeID και αντίστοιχη ακολουθία).

```

<?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);
?>

```

5.4 Υπολογισμός P-value (Bayesian Analysis)

Η p-value αποτελεί ένα μέτρο αξιολόγησης της εγκυρότητας των αποτελεσμάτων που προκύπτουν μέσα από την στατιστική ανάλυση. Προκύπτει από συμπεράσματα στατιστικής φύσεως που δίνουν έμφαση στην συχνότητα εμφάνισης και τις αναλογίες των αποτελεσμάτων. Για να ελέγξουμε την μηδενική υπόθεση, μια γενική δήλωση ή μια προεπιλεγμένη θέση ότι δεν υπάρχει συσχέτιση μεταξύ δύο καταστάσεων, χρειαζόμαστε μια εναλλακτική υπόθεση σε συσχέτισμό με την οποία θα συγκρίνουμε την πρώτη. Η τιμή p-value παίρνει τιμές από το 0 έως το 1. Πιο συγκεκριμένα: μία μικρή τιμή (≤ 0.05) υποδηλώνει ισχυρά πειστήρια ενάντια στην μηδενική υπόθεση, συνεπώς γίνεται αποδεκτή. Αντίστοιχα μία μεγάλη p-value (> 0.05) υποδηλώνει ελλιπή πειστήρια ενάντια στην μηδενική υπόθεση και για αυτό το λόγω απορρίπτεται. Τέλος οι τιμές της p-value πολύ κοντά στο αποδεκτό «επίπεδο λάθους» (κατώφλι ~ 0.05) θεωρούνται οριακές και χρήζουν ιδιαίτερης προσοχής (Zar, 2010). Ένας έλεγχος για την εγκυρότητα των ευρημάτων απαιτείται σε αυτό το στάδιο, λόγω των διαφορών στην πολλαπλότητα τιμών έκφρασης που εμφανίζονται για κάθε γονίδιο (Samuels et al., 2012). Πιο συγκεκριμένα οι υπολογισμοί αυτοί θα διεξαχθούν με σκοπό την σύγκριση των γονιδίων που θα προκύψουν από αυτήν σε σχέση με εκείνα που θα προκύψουν από την μέθοδο της μετα-ανάλυσης.

Τα p-values υπολογίστηκαν μέσω Μπεύζιανής ανάλυσης (Townsend, 2004), που στηρίζεται στην ανάλυση γραμμικής παλινδρόμησης (*Linear Regression Analysis*) και χρησιμοποιεί την κοινή φύση των συμπερασμάτων που εξάγονται από τις μικροσυστοιχίες με σκοπό την εξαγωγή ξεχωριστών συμπερασμάτων για κάθε γονίδιο (Cleveland et al., 1992) μέσα από το περιβάλλον της R (Smyth, 2004). Επειδή ο όγκος δεδομένων που επεξεργάζονται είναι αρκετά μεγάλος επιλέξαμε την χρήση της τροποποιημένης (adjusted) τιμής p-value που προέκυψε με εφαρμογή μεθόδου διόρθωσης πολλαπλών ερωτημάτων FDR (Benjamini and Hochberg, 1995), καθώς μέσω αυτής μειώνεται η πιθανότητα αποδοχής ενός αποτελέσματος ως σπουδαίο λόγω μεγάλης στατιστικής και το οποίο στην πραγματικότητα είναι ψευδώς θετικό. Υπολογίσαμε λοιπόν τις τιμές για κάθε σύγκριση μεταξύ δύο καταστάσεων (control - low και control - high), για τις δύο ομάδες που προέκυψαν με τον διαχωρισμό βάση το χρονικό πλαίσιο μετά την έκθεση, για κάθε γονίδιο έχοντας είδη αντιστοιχίσει κάθε spot με ένα ENSG_ID. Ο κώδικας που χρησιμοποιήθηκε είναι ο εξής:

GSE8917

CONTROL-HIGH

```
>Normalized_Proc = read.table("8917Proc_CH6.txt", header=T)
>agilent.design=model.matrix(~0+factor
/((c(1,2,2,2,1,2,2,2,1,2,2,2,1,2,2,2,1,2,2,2,1,1,1,1,1)))
>colnames(agilent.design) = c("Control", "High")
```

Όπου GSE8917Proc_CH6.txt το αρχείο που περιέχει τις επεξεργασμένες τιμές έκφρασης για κάθε ομάδα (Πίνακας 5). Στην συγκεκριμένη περίπτωση λαμβάνουμε την έκφραση των γονιδίων από την μελέτη GSE8917, από τα δείγματα που ακτινοβολήθηκαν με υψηλή δόση και συλλέχθηκαν μέχρι και 16 ώρες έπειτα από την έκθεση.

Πίνακας 5: Μορφή αρχικού αρχείου ανάγνωσης όπως εισήχθει στο RStudio για τις σταθμισμένες εκφράσεις της σειράς GSE8917 για τον υπολογισμό των p-values

GSE8917	GSM225984	GSM225985	GSM225986	GSM225987	GSM225988
ENSG00000000003	4.21071	4.59369	5.04328	4.93208	3.74542
ENSG00000000005	5.4866	4.91063	4.24242	4.3911	4.46859
ENSG000000000419	10.0681	10.0279	10.0707	9.88983	9.65871
ENSG000000000457	6.57056	7.19536	7.22826	7.15862	6.82968
ENSG000000000460	7.41322	6.30216	6.31345	6.09194	5.62998
ENSG000000000938	7.56387	8.64129	9.14479	9.02293	9.43835
ENSG000000000971	3.92361	4.77827	5.19431	5.09633	5.54313
ENSG00000001036	7.00957	7.39473	7.34622	7.59844	7.30224

Ακολουθούν οι εντολές για την μέθοδο Bayes:

```
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-High,levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results, file = "8917_CH6.txt", quote = F,
/sep = "\t")
```

που εξάγει τον πίνακα σύγκρισης μη ακτινοβολημένων και ακτινοβολημένων με υψηλές δόσεις δειγμάτων που λήφθηκαν μεταξύ 4 και 16 ωρών. Ο *model matrix* ορίζει την κατάσταση για κάθε δείγμα ώστε να χρησιμοποιηθεί από τον πίνακα αντιθέσεων (*contrast matrix*) για την διαδικασία των συγκρίσεων. Ομοίως λειτουργούμε και για την σύγκριση:

CONTROL-LOW

```
>Normalized_Proc = read.table("8917Proc_CL6.txt", header=T)
>agilent.design=model.matrix(~0+factor
/((c(1,2,1,2,1,2,1,2,1,2,1,1,1,1,1)))
>colnames(agilent.design) = c("Control", "Low")
```

```

>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-Low, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results, file = "8917_CL6.txt", quote = F,
/sep = "\t")

```

Τα τελικά αρχεία για κάθε σύγκριση εξήχθησαν από το RStudio και επεξεργάστηκαν για την διατήρηση των απαραίτητων πληροφοριών (Πίνακας 6). Ακολουθεί ο κώδικας για τις υπόλοιπες σειρές δεδομένων.

Πίνακας 6: Ενδεικτική σύγκριση έκφρασης ορισμένων γονιδίων μίας σειράς πειραμάτων που ανήκουν στην ίδια μελέτη και στο ίδιο group ωρών για δύο καταστάσεις χρησιμοποιώντας τις αντίστοιχες τιμές των adjusted p-values.

ENSG00000161513	GSE8917	Control	Low	0.000963
ENSG00000100385	GSE8917	Control	Low	0.001071
ENSG00000127914	GSE8917	Control	Low	0.002179
ENSG00000234062	GSE8917	Control	Low	0.002745
ENSG00000145920	GSE8917	Control	Low	0.002752
ENSG00000129235	GSE8917	Control	Low	0.002827
ENSG00000185615	GSE8917	Control	Low	0.003562
ENSG00000133169	GSE8917	Control	Low	0.003756
ENSG00000276644	GSE8917	Control	Low	0.0044

GSE12435

CONTROL-LOW

```

>Normalized_Proc = read.table("12435Proc_CL4.txt", header=T)
>agilent.design = model.matrix(~ 0+factor (c(2,2,2,2,1,1,1,1)))
>colnames(agilent.design) = c("Control","Low")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-Low, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results,file = "12435_CL4.txt", quote = F,
/sep = "\t")

```

GSE16935

CONTROL-HIGH

```

>Normalized_Proc = read.table("16935Proc_CH16.txt", header=T)
>agilent.design = model.matrix(~ 0+factor ((c(1,1,1,2,2,2,1,1,1))))
>colnames(agilent.design) = c("Control","High")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-High, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)

```

```

>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results,file = "16935_CH16.txt", quote =F,
/sep = "\t")
CONTROL-LOW
>Normalized_Proc = read.table("16935Proc_CL16.txt", header=T)
>agilent.design = model.matrix(~ 0+factor ((c(1,1,1,2,2,2,1,1,1))))
>colnames(agilent.design) = c("Control","Low")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-Low, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results,file = "16935_CL16.txt", quote =F,
/sep = "\t")

```

GSE23901

```

CONTROL-LOW
>Normalized_Proc = read.table("23901Proc_CL16.txt", header=T)
>agilent.design = model.matrix(~ 0+factor(c(1,1,1,2,2,2)))
>colnames(agilent.design) = c("Control","Low")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-Low, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results, file ="23901_CL16.txt", quote = F,
/sep "\t")

```

GSE29344

```

CONTROL-HIGH
>Normalized_Proc = read.table("29344Proc_CH8.txt", header=T)
>agilent.design=model.matrix(~
/0+factor((c(1,1,1,1,1,1,1,1,1,1,2,2))))
>colnames(agilent.design) = c("Control","High")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-High, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Norm))
>write.table(limma.results, file ="29344_CH8.txt", quote = F,
/sep="\t")

```

```

CONTROL-LOW
>Normalized_Proc = read.table("29344Proc_CL8.txt", header=T)
>agilent.design=model.matrix(~
/0+factor((c(1,1,1,1,1,1,1,1,2,2,1,1))))
>colnames(agilent.design) = c("Control","Low")
>limma.fit = lmFit(Normalized_Proc, agilent.design)

```

```

>contrast.matrix=makeContrasts(Control-Low, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results, file = "29344_CL8.txt", quote = F,
/sep = "\t")
GSE59861
CONTROL-HIGH
>Normalized_Proc = read.table("59861Proc_CH6_12.txt", header=T)
>agilent.design = model.matrix(~ 0+factor(rep(1:2,c(4,6))))
>colnames(agilent.design) = c("Control","High")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-High, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results, file = "59861_CH6_12.txt", quote = F,
/sep = "\t")
CONTROL-LOW
>Normalized_Proc = read.table("59861Proc_CL6_12.txt", header=T)
>agilent.design = model.matrix(~ 0+factor(rep(1:2,c(4,6))))
>colnames(agilent.design) = c("Control","Low")
>limma.fit = lmFit(Normalized_Proc, agilent.design)
>contrast.matrix=makeContrasts(Control-Low, levels=agilent.design)
>limma.contrasts.fit = contrasts.fit(limma.fit, contrast.matrix)
>limma.Bayes.fit = eBayes(limma.contrasts.fit)
>limma.results=topTable(limma.Bayes.fit,coef=1,
/number=nrow(Normalized_Proc))
>write.table(limma.results, file = "59861_CL6_12.txt", quote = F,
/sep = "\t")

```

Χρησιμοποιώντας τις προσαρμοσμένες (*adjusted*) τιμές p-values θεωρήσαμε τιμή κατωφλίου 0.05 και επιλέξαμε τα γονίδια που ξεχώρισαν (Script 5). Για τις λίστες γονιδίων που προέκυψαν για κάθε μελέτη σχεδιάσαμε τα αντίστοιχα διαγράμματα Venn ενδεικτικά ώστε να εντοπίσουμε ποσοτικά τα στατιστικώς σημαντικά διαφορικώς εκφρασμένα γονίδια σε κάθε μελέτη αλλά και τα κοινά γονίδια μεταξύ αυτών.

Script 5: p_value.cutOff για επιλογή των στατιστικά σημαντικών διαφορικά εκφρασμένων γονιδίων.

```
<?php
```

```
$low=array("dif.values_12435.Control_Low.txt","dif.values_16935_20629.Control_Low.txt","dif.values_23901.Control_Low.txt","dif.values_29344.Control_Low.txt","dif.values_43151.Control_Low.txt","dif.values_52918.Control_Low.txt","dif.values_59861.Control_Low.txt","dif.values_8917.Control_Low.txt");
```

```
$high=array("dif.values_59861.Control_High.txt","dif.values_8917.Control_High.txt","dif.values_16935_20629.Control_High.txt");
```

```
calc($low);calc($high);
```

```
function calc($marray) {  
    $marraynum=count($marray);  
    //echo $marraynum;  
    for($i=0;$i<$marraynum;$i++){  
        $venn=$marray[$i];  
        $venn=str_replace(".txt",".ven",$venn);  
        //echo "$marray[$i]\t$venn\n";  
        $fh=fopen($venn,"w");  
        $line=file($marray[$i]);  
        $linenum=count($line);  
        //echo "$linenum\n";  
        for($j=1;$j<$linenum;$j++) {  
            $line[$j]=rtrim($line[$j]);  
            $arr=explode("\t",$line[$j]);  
            if($arr[5]<=0.1) {  
                fwrite($fh,"$arr[0]\n");  
            }  
        }  
        fclose($fh);  
    }  
}
```

```
?>
```

5.5 Μετα-ανάλυση (STATA)

Μετά την εξαγωγή των σταθμισμένων εκφράσεων για κάθε γονίδιο από την βάση δεδομένων έγινε ο διαχωρισμός των δειγμάτων με βάση το χρονικό πλαίσιο (Script 6). Η εφαρμογή της μεθόδου μετα-ανάλυσης έγινε στο στατιστικό πρόγραμμα STATA (StataCorp, 2013). Καμία μέθοδος διόρθωσης πολλαπλών ερωτημάτων δεν ανέδειξε κάποιο στατιστικά σημαντικό διαφορικά εκφραζόμενο γονίδιο. Επιλέξαμε λοιπόν μία πιο αυστηρή τιμή κατωφλίου (0.01) για το μη τροποποιημένο p-value και καταλήξαμε σε ορισμένα στατιστικώς σημαντικά διαφορικώς εκφραζόμενα γονίδια (Πίνακας 8, Πίνακας 9). Συμπεριλήφθηκε το όνομα στο HGNC και μια σύντομη περιγραφή κάθε γονιδίου από το BioMart.

Script 6: Ionising_Hours.php για τον διαχωρισμό των δειγμάτων με βάση την ώρα που συλλέχθηκαν.

```
<?php
$line=file("GSM.txt");
$linenum=count($line);
for($i=0;$i<$linenum;$i++){
    $line[$i]=rtrim($line[$i]);
    $arr=explode("\t",$line[$i]);
    $GSM[$arr[0]]=$arr[1];
}
//print_r($GSM);
$n=$argv[1];
$line=file("Ionising.txt");
$linenum=count($line);
echo $line[0];
for($i=1;$i<$linenum;$i++){
    if(strncmp($line[$i],"GSE43151",8)) {
        $arr=explode("\t",$line[$i]);
        //echo "$arr[1]\n";
        if(!strcmp($arr[1],"Control")) {
            echo $line[$i];
        }
        else {
            if($GSM[$arr[2]]==$n) {
                echo $line[$i];
            }
        }
        //echo $line[$i];
    }
}
?>
```

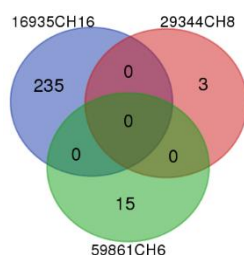

5.6 Βιβλιογραφική Αναζήτηση (Quertle-Cosmic)

Χάριν πληρότητας αυτής της διπλωματικής εργασίας διεξήχθει συμπληρωματικά μια βιβλιογραφική αναζήτηση στο Quertle (Giglia, 2011), για μελέτες ή πειράματα, σχετικά με την έκθεση σε ιονίζουσες ακτινοβολίες κυρίως χαμηλών δόσεων, στα οποία δεν χρησιμοποιήθηκαν δεδομένα μικροσυστοιχιών ή αυτά δεν συμπεριλαμβάνονταν στο GEO . Σκοπός είναι να εντοπιστούν τα διαφορεικά εκφραζόμενα γονίδια στις μελέτες αυτές (Πίνακας 33, Πίνακας 34) και να γίνει μια σύγκριση με τα δικά μας αποτελέσματα. Χρησιμοποιώντας τους «έξυπνους» όρους αναζήτησης του Quertle (*power terms*) εκτελέσαμε την αναζήτηση: `$protein low dose radiation human`. Ο διαχωρισμός των εν λόγω γονιδίων έγινε δια χειρός. Για να είναι δυνατή η σύγκριση των ευρημάτων από τις μελέτες γονιδιακής απόκρισης που βρέθηκαν (Long et al., 2007; Stankevicius et al., 2013; Wang et al., 2006) αντιστοιχίσαμε την ονοματολογία που χρησιμοποίησαν με το τελευταίο όνομα στο HGNC για κάθε γονίδιο, αναζητώντας το πρωταρχικό γονίδιο ή το μετάγραφο του στο GeneCards (Safran et al., 2010) και το GenBank (Benson et al., 2013) ή χρησιμοποιώντας την εφαρμογή BLAT (Kent, 2002) στο Ensembl (Cunningham et al., 2015). Επιπλέον χρησιμοποιώντας την λίστα *cancer gene sensus* (Futreal et al., 2004) την οποία κατεβάσαμε από τον κατάλογο COSMIC (Forbes et al., 2008) σχεδιάστηκε το ανάλογο διάγραμμα Venn (Εικόνα 39) με τα γονίδια που προέκυψαν από αυτή την ανάλυση και σε συνδυασμό με την βιβλιογραφία.

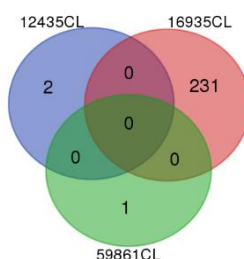
6 ΑΠΟΤΕΛΕΣΜΑΤΑ

6.1 Αρχικός Έλεγχος με Διαγράμματα Venn

Χρησιμοποιώντας τις τροποποιημένες τιμές p -value με μέθοδο διόρθωσης πολλαπλών ερωτημάτων FDR, ορισμένες μελέτες δεν έδειξαν κανένα στατιστικώς σημαντικό διαφορικώς εκφραζόμενο γονίδιο. Για τις υψηλές και για τις χαμηλές δόσεις δεν βρέθηκε ένα στατιστικώς σημαντικό γονίδιο κοινό σε πάνω από μία μελέτη (Εικόνα 28, Εικόνα 29). Παρόλα αυτά επιβεβαιώθηκαν τα προβλεπόμενα αποτελέσματα όπου αναμένονταν περισσότερα διαφορικώς εκφρασμένα γονίδια έπειτα από έκθεση σε υψηλές δόσεις σε σχέση με τις χαμηλές. Η εικόνα των διαγραμμάτων Venn αποτελεί μια πρώτη ένδειξη για την πιθανή λανθασμένη επιλογή και διαχείριση των μελετών αυτής της ανάλυσης. Παρόλα αυτά δίνοντας βαρύτητα στο μοντέλο των τυχαίων επιδράσεων, ακολούθησε η διεξαγωγή μετα-ανάλυσης.



Εικόνα 28: Διαγράμματα Venn με cutoff=0.05 για το p -value, για δείγματα από έκθεση σε υψηλή δόση ακτινοβολίας με κριτήριο το adjusted p -value από Μπεϋζιανή ανάλυση.



Εικόνα 29: Διαγράμματα Venn με cutoff=0.05 για το p -value, για δείγματα από έκθεση σε χαμηλή δόση ακτινοβολίας με κριτήριο το adjusted p -value από Μπεϋζιανή ανάλυση.

6.2 Μετα-ανάλυση

Ο αριθμός γονιδίων που προέκυψαν (Πίνακας 7) θέτοντας όριο για το μη τροποποιημένο p-value την τιμή 0.01, είναι αρκετά μεγάλος εφόσον δεν λάβαμε υπόψιν τις μεθόδους διόρθωσης πολλαπλών ερωτημάτων, οι οποίες δεν ανέδειξαν κανένα στατιστικώς σημαντικό διαφορικά εκφραζόμενο γονίδιο. Συμπεραίνουμε λοιπόν πως ορισμένα από τα γονίδια αυτά είναι ψευδώς θετικά και άρα στην πραγματικότητα δεν είναι στατιστικώς σημαντικά. Έτσι θα επιδιώξουμε με την βοήθεια δύο εργαλείων βιοπληροφορικής να τα ταξινομήσουμε και να τα ομαδοποιήσουμε. Σκοπός είναι να ξεχωρίσουμε τα σημαντικότερα με βάση την βιολογική τους λειτουργία και τις συνδέσεις που προκύπτουν μεταξύ αυτών.

Πίνακας 7: Αριθμός γονιδίων που αναδείχθηκαν ως στατιστικώς σημαντικά από την μετα-ανάλυση βάση της τιμής p-value<0.01 για κάθε σύνολο γονιδίων.

Group	Control-High 4-16	Control Low 4-16
#Genes	571	266

• Τελικοί Πίνακες Διαφορικώς Εκφρασμένων Γονιδίων

Πίνακας 8: Γονίδια που ξεχώρισαν ως διαφορικά εκφρασμένα μέσω μετα-ανάλυσης, με βάση τιμή p-value<0.01, έπειτα από έκθεση σε υψηλή δόση ακτινοβολίας(>0.5Gy) στο διάστημα 4-16 ωρών. Συμπεριλαμβάνεται το αντίστοιχο HGNC σύμβολο με την αντίστοιχη περιγραφή και οι ακριβείς τιμές του υπολογισμένου p-value και της διαφοράς στην έκφραση.

ENSG	HGNC	Description	Diff	P-value
ENSG00000272223			-1.76909	0.000013
ENSG00000263324	HCN3	hyperpolarization activated cyclic nucleotide gated potassium channel 3 [Source:HGNC Symbol;Acc:HGNC:19183]	-1.43019	0.000041
ENSG00000204437	CTSLP6	cathepsin L pseudogene 6 [Source:HGNC Symbol;Acc:HGNC:23611]	-1.58584	0.000053
ENSG00000230011	CTSLP4	cathepsin L pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:23645]	-1.58584	0.000053
ENSG00000236417	CTSLP1	cathepsin L pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:2539]	-1.58584	0.000053
ENSG00000131080	EDA2R	ectodysplasin A2 receptor [Source:HGNC Symbol;Acc:HGNC:17756]	-1.37113	0.000094
ENSG00000187997	C17orf99	chromosome 17 open reading frame 99 [Source:HGNC Symbol;Acc:HGNC:34490]	-1.48835	0.000126
ENSG00000272081			-1.42651	0.000205
ENSG00000185885	IFITM1	interferon induced transmembrane protein 1 [Source:HGNC Symbol;Acc:HGNC:5412]	1.28088	0.000235
ENSG00000059769	DNAJC25	DnaJ heat shock protein family (Hsp40) member C25 [Source:HGNC Symbol;Acc:HGNC:34187]	-1.25641	0.000247
ENSG00000275430	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.24624	0.000252
ENSG00000259188			-1.38678	0.000295
ENSG00000238098	ABCA17P	ATP binding cassette subfamily A member 17, pseudogene [Source:HGNC Symbol;Acc:HGNC:32972]	-1.38393	0.000296
ENSG00000273024	INTS4P2	integrator complex subunit 4 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:22351]	-1.37911	0.000299

ENSG00000106479	ZNF862	zinc finger protein 862 [Source:HGNC Symbol;Acc:HGNC:34519]	-1.37254	0.000317
ENSG00000226995	LINC00658	long intergenic non-protein coding RNA 658 [Source:HGNC Symbol;Acc:HGNC:44315]	-1.37205	0.000324
ENSG00000263293		thyroid cancer-associated transcript 158 [Source:EntrezGene;Acc:102724508]	-1.36724	0.000339
ENSG00000224507	HCG23	HLA complex group 23 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:19713]	-1.36318	0.000348
ENSG00000225834	HCG23	HLA complex group 23 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:19713]	-1.36318	0.000348
ENSG00000226228	HCG23	HLA complex group 23 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:19713]	-1.36318	0.000348
ENSG00000229903	HCG23	HLA complex group 23 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:19713]	-1.36318	0.000348
ENSG00000230602	HCG23	HLA complex group 23 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:19713]	-1.36318	0.000348
ENSG00000262484	CCER2	coiled-coil glutamate rich protein 2 [Source:HGNC Symbol;Acc:HGNC:44662]	-1.35915	0.000357
ENSG00000283099	CCER2	coiled-coil glutamate rich protein 2 [Source:HGNC Symbol;Acc:HGNC:44662]	-1.35915	0.000357
ENSG00000263826			-1.35956	0.000361
ENSG00000231528	FAM225A	family with sequence similarity 225 member A (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:27855]	-1.36054	0.000362
ENSG00000273391			-1.35488	0.000368
ENSG00000270761			-1.35359	0.000371
ENSG00000228192			-1.35083	0.000383
ENSG00000169118	CSNK1G1	casein kinase 1 gamma 1 [Source:HGNC Symbol;Acc:HGNC:2454]	-1.24359	0.000384
ENSG00000244280	ECEL1P2	endothelin converting enzyme like 1 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:14019]	-1.34753	0.000399
ENSG00000262943	ALOX12P2	arachidonate 12-lipoxygenase pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:432]	-1.21925	0.000403
ENSG00000275963			-1.34124	0.000416
ENSG00000120262	CCDC170	coiled-coil domain containing 170 [Source:HGNC Symbol;Acc:HGNC:21177]	-1.34104	0.000418
ENSG00000260314	MRC1	mannose receptor C-type 1 [Source:HGNC Symbol;Acc:HGNC:7228]	-1.33655	0.000428
ENSG00000189152	GRAPL	GRB2 related adaptor protein like [Source:HGNC Symbol;Acc:HGNC:37240]	-1.33113	0.000459
ENSG00000260223			-1.33072	0.000459
ENSG00000248699			-1.32709	0.000469
ENSG00000281205	LINC00950	long intergenic non-protein coding RNA 950 [Source:HGNC Symbol;Acc:HGNC:28098]	-1.32432	0.000475
ENSG00000229891	LINC01315	long intergenic non-protein coding RNA 1315 [Source:HGNC Symbol;Acc:HGNC:50513]	-1.32393	0.000479
ENSG00000227456	LINC00310	long intergenic non-protein coding RNA 310 [Source:HGNC Symbol;Acc:HGNC:16414]	-1.32604	0.000482
ENSG00000256577			-1.32119	0.000494
ENSG00000280202			-1.31951	0.0005
ENSG00000276131			-1.31669	0.000505
ENSG00000235194	PPP1R3E	protein phosphatase 1 regulatory subunit 3E [Source:HGNC Symbol;Acc:HGNC:14943]	-1.31621	0.000509
ENSG00000223750	SIRPB3P	signal regulatory protein beta 3, pseudogene [Source:HGNC Symbol;Acc:HGNC:49209]	-1.31499	0.000519
ENSG00000245532	NEAT1	nuclear paraspeckle assembly transcript 1 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:30815]	-1.31289	0.00052
ENSG00000187486	KCNJ11	potassium voltage-gated channel subfamily J member 11 [Source:HGNC Symbol;Acc:HGNC:6257]	-1.31655	0.000527
ENSG00000269845			-1.29993	0.00059

ENSG00000279112	BRI3BPP1	BRI3 binding protein pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:51482]	-1.29993	0.00059
ENSG00000269994			-1.28931	0.000637
ENSG00000239533	GOLGA2P2Y	golgin A2 pseudogene 2, Y-linked [Source:HGNC Symbol;Acc:HGNC:16479]	-1.15253	0.000647
ENSG00000213529			-1.27828	0.000696
ENSG00000080546	SESN1	sestrin 1 [Source:HGNC Symbol;Acc:HGNC:21595]	-1.27434	0.000748
ENSG00000166763	STRCP1	stereocilin pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:33915]	-1.26988	0.000753
ENSG00000223489	NEFHP1	neurofilament heavy pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:7738]	-1.26674	0.000768
ENSG00000181638	ZFP41	ZFP41 zinc finger protein [Source:HGNC Symbol;Acc:HGNC:26786]	-1.1364	0.000774
ENSG00000224916	APOC4- APOC2	APOC4-APOC2 readthrough (NMD candidate) [Source:HGNC Symbol;Acc:HGNC:44426]	-1.14322	0.000781
ENSG00000232286			-1.26506	0.000783
ENSG00000124762	CDKN1A	cyclin dependent kinase inhibitor 1A [Source:HGNC Symbol;Acc:HGNC:1784]	-1.35316	0.000788
ENSG00000260565	ERVK13-1	endogenous retrovirus group K13 member 1 [Source:HGNC Symbol;Acc:HGNC:27548]	-1.12917	0.000799
ENSG00000078237	TIGAR	TP53 induced glycolysis regulatory phosphatase [Source:HGNC Symbol;Acc:HGNC:1185]	-1.07478	0.00082
ENSG00000139269	INHBE	inhibin beta E subunit [Source:HGNC Symbol;Acc:HGNC:24029]	-1.03738	0.000826
ENSG00000218739	CEBPZOS	CEBPZ opposite strand [Source:HGNC Symbol;Acc:HGNC:49288]	-1.2567	0.00083
ENSG00000236581	STARD13-AS	STARD13 antisense RNA [Source:HGNC Symbol;Acc:HGNC:40873]	-1.25694	0.000831
ENSG00000254926			-1.25622	0.000839
ENSG00000229391	HLA-DRB6	major histocompatibility complex, class II, DR beta 6 (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4954]	-1.12278	0.000847
ENSG00000216649	GAGE12E	G antigen 12E [Source:HGNC Symbol;Acc:HGNC:31905]	-3.52997	0.000848
ENSG00000185088	RPS27L	ribosomal protein S27 like [Source:HGNC Symbol;Acc:HGNC:18476]	-1.01265	0.000858
ENSG00000237188			-1.25598	0.00088
ENSG00000273071			-1.25598	0.00088
ENSG00000176945	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000274843	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000275501	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000276583	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000277753	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000278114	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000281630	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.10161	0.001035
ENSG00000279785			-1.23026	0.001039
ENSG00000268996	MAN1B1-AS1	MAN1B1 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48715]	-1.22518	0.001067
ENSG00000112200	ZNF451	zinc finger protein 451 [Source:HGNC Symbol;Acc:HGNC:21091]	-0.98523	0.001081
ENSG00000242086	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000280512	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000280993	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000281060	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088

ENSG00000281603	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000281794	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000281915	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000282953	LINC00969	long intergenic non-protein coding RNA 969 [Source:HGNC Symbol;Acc:HGNC:48729]	-1.09553	0.001088
ENSG00000225968	ELFN1	extracellular leucine rich repeat and fibronectin type III domain containing 1 [Source:HGNC Symbol;Acc:HGNC:33154]	-1.22143	0.001106
ENSG00000101323	HAO1	hydroxyacid oxidase 1 [Source:HGNC Symbol;Acc:HGNC:4809]	-0.98311	0.001113
ENSG00000110244	APOA4	apolipoprotein A4 [Source:HGNC Symbol;Acc:HGNC:602]	-0.97888	0.001133
ENSG00000196152	ZNF79	zinc finger protein 79 [Source:HGNC Symbol;Acc:HGNC:13153]	-0.97047	0.001161
ENSG00000174943	KCTD13	potassium channel tetramerization domain containing 13 [Source:HGNC Symbol;Acc:HGNC:22234]	-0.97389	0.001175
ENSG00000149548	CCDC15	coiled-coil domain containing 15 [Source:HGNC Symbol;Acc:HGNC:25798]	-1.21271	0.001185
ENSG00000259780			-1.21089	0.001199
ENSG00000164669	INTS4P1	integrator complex subunit 4 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:21925]	-1.20876	0.001218
ENSG00000106125	MINDY4	MINDY lysine 48 deubiquitinase 4 [Source:HGNC Symbol;Acc:HGNC:21916]	-1.07481	0.001319
ENSG00000249679			-1.19861	0.001322
ENSG00000172508	CARNS1	carnosine synthase 1 [Source:HGNC Symbol;Acc:HGNC:29268]	-1.19913	0.001323
ENSG00000237872	POU5F1P4	POU class 5 homeobox 1 pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:33310]	-1.19878	0.001325
ENSG00000235602	POU5F1P3	POU class 5 homeobox 1 pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:9222]	-1.1987	0.001326
ENSG00000282901			3.2904	0.00133
ENSG00000206478	IER3	immediate early response 3 [Source:HGNC Symbol;Acc:HGNC:5392]	-3.28027	0.001357
ENSG00000227942	FRMD8P1	FERM domain containing 8 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:24690]	-1.19441	0.001367
ENSG00000211626	IGKV6D-41	immunoglobulin kappa variable 6D-41 (non-functional) [Source:HGNC Symbol;Acc:HGNC:5838]	-1.19334	0.001385
ENSG00000261335			-1.1937	0.001405
ENSG00000276111	SDCCAG8	serologically defined colon cancer antigen 8 [Source:HGNC Symbol;Acc:HGNC:10671]	-1.08172	0.001476
ENSG00000214331			-3.23432	0.001484
ENSG00000261897	OR5T3	olfactory receptor family 5 subfamily T member 3 [Source:HGNC Symbol;Acc:HGNC:15297]	-3.22796	0.001502
ENSG00000144648	ACKR2	atypical chemokine receptor 2 [Source:HGNC Symbol;Acc:HGNC:1565]	-1.06595	0.001507
ENSG00000135253	KCP	kielin/chordin-like protein [Source:HGNC Symbol;Acc:HGNC:17585]	-1.18257	0.001508
ENSG00000204188	GGNBP1	gametogenetin binding protein 1 (pseudogene) [Source:HGNC Symbol;Acc:HGNC:19427]	-1.18025	0.001549
ENSG00000228102	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158
ENSG00000228912	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158
ENSG00000229942	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158
ENSG00000230722	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158

ENSG00000233023	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158
ENSG00000233298	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158
ENSG00000235821	IFITM4P	interferon induced transmembrane protein 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:21669]	1.06753	0.00158
ENSG00000236914	LINC01852	long intergenic non-protein coding RNA 1852 [Source:HGNC Symbol;Acc:HGNC:52668]	-1.17603	0.001584
ENSG00000239521	GATS	GATS, stromal antigen 3 opposite strand [Source:HGNC Symbol;Acc:HGNC:29954]	-1.17437	0.001618
ENSG00000170983	LINC00208	long intergenic non-protein coding RNA 208 [Source:HGNC Symbol;Acc:HGNC:15535]	-1.17306	0.001633
ENSG00000163206	SMCP	sperm mitochondria associated cysteine rich protein [Source:HGNC Symbol;Acc:HGNC:6962]	-0.94948	0.001635
ENSG00000237797			-1.17562	0.00164
ENSG00000114019	AMOTL2	angiomin like 2 [Source:HGNC Symbol;Acc:HGNC:17812]	-1.17133	0.001681
ENSG00000221990	EXOC3-AS1	EXOC3 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:25175]	-1.16735	0.001699
ENSG00000273722	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000274384	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000274873	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000274905	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000276260	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000277667	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000277789	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000277996	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000278363	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-1.05572	0.001712
ENSG00000240005			-1.16648	0.00172
ENSG00000278238			-1.16568	0.001721
ENSG00000171467	ZNF318	zinc finger protein 318 [Source:HGNC Symbol;Acc:HGNC:13578]	-0.94187	0.001749
ENSG00000158639	PAGE5	PAGE family member 5 [Source:HGNC Symbol;Acc:HGNC:29992]	-0.93517	0.001768
ENSG00000124406	ATP8A1	ATPase phospholipid transporting 8A1 [Source:HGNC Symbol;Acc:HGNC:13531]	-0.95629	0.001778
ENSG00000176937			-3.12777	0.001834
ENSG00000279270	OR52R1	olfactory receptor family 52 subfamily R member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:15235]	-3.12777	0.001834
ENSG00000256614	AK6P1	adenylate kinase 6 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:51678]	-1.15705	0.001844
ENSG00000164556	FAM183BP	family with sequence similarity 183 member B, pseudogene [Source:HGNC Symbol;Acc:HGNC:34511]	-1.15645	0.001861
ENSG00000182993	C12orf60	chromosome 12 open reading frame 60 [Source:HGNC Symbol;Acc:HGNC:28726]	-0.95672	0.001895
ENSG00000224446	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000224582	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000225426	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193

ENSG00000226687	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000229383	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000230674	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000233857	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000235788	LINC01015	long intergenic non-protein coding RNA 1015 [Source:HGNC Symbol;Acc:HGNC:48988]	-1.15195	0.00193
ENSG00000166321	NUDT13	nudix hydrolase 13 [Source:HGNC Symbol;Acc:HGNC:18827]	-0.93051	0.001942
ENSG00000236155			-1.15075	0.001944
ENSG00000150991	UBC	ubiquitin C [Source:HGNC Symbol;Acc:HGNC:12468]	-1.00101	0.001961
ENSG00000260391			-1.1478	0.001991
ENSG00000135047	CTSL	cathepsin L [Source:HGNC Symbol;Acc:HGNC:2537]	-1.02337	0.002007
ENSG00000198755	RPL10A	ribosomal protein L10a [Source:HGNC Symbol;Acc:HGNC:10299]	-1.02684	0.002054
ENSG00000255776	VDAC2P2	voltage dependent anion channel 2 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:32693]	-1.14335	0.00206
ENSG00000167553	TUBA1C	tubulin alpha 1c [Source:HGNC Symbol;Acc:HGNC:20768]	-1.11573	0.002101
ENSG00000197632	SERPINB2	serpin family B member 2 [Source:HGNC Symbol;Acc:HGNC:8584]	-1.03159	0.002141
ENSG00000172244	C5orf34	chromosome 5 open reading frame 34 [Source:HGNC Symbol;Acc:HGNC:24738]	-1.02392	0.002175
ENSG00000102606	ARHGEF7	Rho guanine nucleotide exchange factor 7 [Source:HGNC Symbol;Acc:HGNC:15607]	-0.91663	0.002204
ENSG00000103042	SLC38A7	solute carrier family 38 member 7 [Source:HGNC Symbol;Acc:HGNC:25582]	-0.91825	0.002213
ENSG00000249493	ANKRD20A18P	ankyrin repeat domain 20 family member A18, pseudogene [Source:HGNC Symbol;Acc:HGNC:23756]	-1.13286	0.002235
ENSG00000256315			-1.13124	0.002266
ENSG00000258603			-1.13125	0.002282
ENSG00000272017			-1.23882	0.002286
ENSG00000245149	RNF139-AS1	RNF139 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48940]	-1.13061	0.002294
ENSG00000101230	ISM1	isthmin 1 [Source:HGNC Symbol;Acc:HGNC:16213]	-1.12852	0.002312
ENSG00000229694		long intergenic non-protein coding RNA 484 [Source:EntrezGene;Acc:100129347]	-1.12826	0.002324
ENSG00000188004	C1orf204	chromosome 1 open reading frame 204 [Source:HGNC Symbol;Acc:HGNC:27647]	-1.12603	0.002387
ENSG00000183943	PRKX	protein kinase, X-linked [Source:HGNC Symbol;Acc:HGNC:9441]	-1.1248	0.002389
ENSG00000226810			-1.12334	0.002411
ENSG00000173401	GLIPR1L1	GLI pathogenesis related 1 like 1 [Source:HGNC Symbol;Acc:HGNC:28392]	-0.91004	0.002418
ENSG00000143434	SEMA6C	semaphorin 6C [Source:HGNC Symbol;Acc:HGNC:10740]	-0.90706	0.002442
ENSG00000119777	TMEM214	transmembrane protein 214 [Source:HGNC Symbol;Acc:HGNC:25983]	-1.13075	0.002454
ENSG00000248327	NOL8P1	nucleolar protein 8 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:49557]	-1.12069	0.002473
ENSG00000279672			-1.11917	0.002489
ENSG00000197813			-1.119	0.002491
ENSG00000135423	GLS2	glutaminase 2 [Source:HGNC Symbol;Acc:HGNC:29570]	-0.90467	0.002501

ENSG00000243701	DUBR	DPPA2 upstream binding RNA [Source:HGNC Symbol;Acc:HGNC:48569]	-1.11788	0.002519
ENSG00000229021			-1.11752	0.002522
ENSG00000240667			-1.11752	0.002522
ENSG00000185532	PRKG1	protein kinase, cGMP-dependent, type I [Source:HGNC Symbol;Acc:HGNC:9414]	-1.11695	0.002531
ENSG00000205269	TMEM170B	transmembrane protein 170B [Source:HGNC Symbol;Acc:HGNC:34244]	-1.11625	0.002546
ENSG00000211789	TRAV12-2	T-cell receptor alpha variable 12-2 [Source:HGNC Symbol;Acc:HGNC:12106]	-1.11551	0.002569
ENSG00000144840	RABL3	RAB, member of RAS oncogene family like 3 [Source:HGNC Symbol;Acc:HGNC:18072]	-0.90354	0.002591
ENSG00000102100	SLC35A2	solute carrier family 35 member A2 [Source:HGNC Symbol;Acc:HGNC:11022]	-0.90054	0.002606
ENSG00000259890	DNM1P50	dynamamin 1 pseudogene 50 [Source:HGNC Symbol;Acc:HGNC:48499]	-1.11315	0.002609
ENSG00000261708	DNM1P32	dynamamin 1 pseudogene 32 [Source:HGNC Symbol;Acc:HGNC:35179]	-1.11315	0.002609
ENSG00000261792	DNM1P28	dynamamin 1 pseudogene 28 [Source:HGNC Symbol;Acc:HGNC:35175]	-1.11315	0.002609
ENSG00000274966			-1.11315	0.002609
ENSG00000282324	DNM1P28	dynamamin 1 pseudogene 28 [Source:HGNC Symbol;Acc:HGNC:35175]	-1.11315	0.002609
ENSG00000282701	DNM1P50	dynamamin 1 pseudogene 50 [Source:HGNC Symbol;Acc:HGNC:48499]	-1.11315	0.002609
ENSG00000228463			-1.11332	0.002633
ENSG00000225469			-1.11228	0.002639
ENSG00000255389			-1.11144	0.002643
ENSG00000131951	LRRC9	leucine rich repeat containing 9 [Source:HGNC Symbol;Acc:HGNC:19848]	-1.11071	0.002664
ENSG00000280401			-1.11	0.002676
ENSG00000197171	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000224633	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000227178	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000231279	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000232146	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000232469	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000237690	OR2B4P	olfactory receptor family 2 subfamily B member 4 pseudogene [Source:HGNC Symbol;Acc:HGNC:8239]	-1.10987	0.002677
ENSG00000170379	TCAF2	TRPM8 channel associated factor 2 [Source:HGNC Symbol;Acc:HGNC:26878]	-0.99718	0.002723
ENSG00000099399	MAGEB2	MAGE family member B2 [Source:HGNC Symbol;Acc:HGNC:6809]	-1.10698	0.002741
ENSG00000104450	SPAG1	sperm associated antigen 1 [Source:HGNC Symbol;Acc:HGNC:11212]	-0.89496	0.002785
ENSG00000236445	LINC00608	long intergenic non-protein coding RNA 608 [Source:HGNC Symbol;Acc:HGNC:27179]	-1.10388	0.002814
ENSG00000259660	DNM1P47	dynamamin 1 pseudogene 47 [Source:HGNC Symbol;Acc:HGNC:35200]	-1.10285	0.002826
ENSG00000261096			-1.10221	0.00284
ENSG00000142798	HSPG2	heparan sulfate proteoglycan 2 [Source:HGNC Symbol;Acc:HGNC:5273]	-0.8906	0.002854
ENSG00000174327	SLC16A13	solute carrier family 16 member 13 [Source:HGNC Symbol;Acc:HGNC:31037]	-0.89156	0.002869

ENSG00000232850	PTGES2-AS1	PTGES2 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48711]	-1.0998	0.002901
ENSG00000170448	NFXL1	nuclear transcription factor, X-box binding like 1 [Source:HGNC Symbol;Acc:HGNC:18726]	-0.88916	0.00299
ENSG00000236144	TMEM147-AS1	TMEM147 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:51273]	-1.09455	0.003015
ENSG00000214720	KRT18P49	keratin 18 pseudogene 49 [Source:HGNC Symbol;Acc:HGNC:33419]	-1.09581	0.00302
ENSG00000272477			-1.09395	0.003035
ENSG00000112245	PTP4A1	protein tyrosine phosphatase type IVA, member 1 [Source:HGNC Symbol;Acc:HGNC:9634]	-0.87778	0.003086
ENSG00000262001	DLGAP1-AS2	DLGAP1 antisense RNA 2 [Source:HGNC Symbol;Acc:HGNC:28146]	-1.09025	0.003121
ENSG00000269343	ZNF587B	zinc finger protein 587B [Source:HGNC Symbol;Acc:HGNC:37142]	-1.09043	0.003125
ENSG00000246731		uncharacterized protein MGC16275 [Source:EntrezGene;Acc:85001]	-1.08769	0.0032
ENSG00000260339	HEXA-AS1	HEXA antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:25810]	-1.08652	0.003211
ENSG00000143847	PPFIA4	PTPRF interacting protein alpha 4 [Source:HGNC Symbol;Acc:HGNC:9248]	-0.87661	0.003262
ENSG00000251193			-1.08456	0.003265
ENSG00000156453	PCDH1	protocadherin 1 [Source:HGNC Symbol;Acc:HGNC:8655]	-0.87774	0.003296
ENSG00000155850	SLC26A2	solute carrier family 26 member 2 [Source:HGNC Symbol;Acc:HGNC:10994]	-1.00001	0.003341
ENSG00000269720			-1.08109	0.003348
ENSG00000228434			-1.07917	0.003399
ENSG00000243155			-1.07547	0.003504
ENSG00000268883	PNMA6B	paraneoplastic Ma antigen family member 6B (pseudogene) [Source:HGNC Symbol;Acc:HGNC:26681]	-1.07421	0.003534
ENSG00000175470	PPP2R2D	protein phosphatase 2 regulatory subunit Bdelta [Source:HGNC Symbol;Acc:HGNC:23732]	-0.86804	0.003547
ENSG00000278633			-1.07338	0.003555
ENSG00000055955	ITIH4	inter-alpha-trypsin inhibitor heavy chain family member 4 [Source:HGNC Symbol;Acc:HGNC:6169]	-0.96655	0.003582
ENSG00000260372	AQP4-AS1	AQP4 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:26399]	-1.07067	0.00363
ENSG00000279484	KLHL30-AS1	KLHL30 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:31018]	-1.06976	0.003657
ENSG00000214223	HNRNPA1P10	heterogeneous nuclear ribonucleoprotein A1 pseudogene 10 [Source:HGNC Symbol;Acc:HGNC:39128]	-2.79362	0.00368
ENSG00000236383	LINC00854	long intergenic non-protein coding RNA 854 [Source:HGNC Symbol;Acc:HGNC:43658]	-1.06883	0.003681
ENSG00000237265			-1.06788	0.00371
ENSG00000049167	ERCC8	ERCC excision repair 8, CSA ubiquitin ligase complex subunit [Source:HGNC Symbol;Acc:HGNC:3439]	-0.8629	0.003718
ENSG00000170837	GPR27	G protein-coupled receptor 27 [Source:HGNC Symbol;Acc:HGNC:4482]	-0.86307	0.003722
ENSG00000228705	LINC00659	long intergenic non-protein coding RNA 659 [Source:HGNC Symbol;Acc:HGNC:44316]	-1.06708	0.003732
ENSG00000263012	KRT33B	keratin 33B [Source:HGNC Symbol;Acc:HGNC:6451]	-0.97088	0.003734
ENSG00000254533			-1.06699	0.00374
ENSG00000282524			-1.06699	0.00374
ENSG00000227543	SPAG5-AS1	SPAG5 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:41140]	-1.06662	0.00375
ENSG00000112394	SLC16A10	solute carrier family 16 member 10 [Source:HGNC Symbol;Acc:HGNC:17027]	-0.86408	0.003766

ENSG00000187607	ZNF286A	zinc finger protein 286A [Source:HGNC Symbol;Acc:HGNC:13501]	-0.96347	0.003805
ENSG00000204745			-1.14942	0.003849
ENSG00000170855	TRIAP1	TP53 regulated inhibitor of apoptosis 1 [Source:HGNC Symbol;Acc:HGNC:26937]	-0.85569	0.003873
ENSG00000113328	CCNG1	cyclin G1 [Source:HGNC Symbol;Acc:HGNC:1592]	-0.86819	0.003892
ENSG00000176974	SHMT1	serine hydroxymethyltransferase 1 [Source:HGNC Symbol;Acc:HGNC:10850]	-0.86057	0.003912
ENSG00000169618	PROKR1	prokineticin receptor 1 [Source:HGNC Symbol;Acc:HGNC:4524]	-0.96463	0.003927
ENSG00000166863	TAC3	tachykinin 3 [Source:HGNC Symbol;Acc:HGNC:11521]	-1.02682	0.003935
ENSG00000176422	SPRYD4	SPRY domain containing 4 [Source:HGNC Symbol;Acc:HGNC:27468]	-0.85977	0.003949
ENSG00000259316			-0.96563	0.00395
ENSG00000280439			-1.05865	0.003982
ENSG00000184060	ADAP2	ArfGAP with dual PH domains 2 [Source:HGNC Symbol;Acc:HGNC:16487]	-0.85568	0.003999
ENSG00000130684	ZNF337	zinc finger protein 337 [Source:HGNC Symbol;Acc:HGNC:15809]	-0.85979	0.004022
ENSG00000142686	C1orf216	chromosome 1 open reading frame 216 [Source:HGNC Symbol;Acc:HGNC:26800]	-0.85823	0.00407
ENSG00000122877	EGR2	early growth response 2 [Source:HGNC Symbol;Acc:HGNC:3239]	-1.0563	0.004079
ENSG00000180712			-1.05419	0.004123
ENSG00000182853	VMO1	vitelline membrane outer layer 1 homolog [Source:HGNC Symbol;Acc:HGNC:30387]	-1.1535	0.004136
ENSG00000184566			-1.05239	0.004176
ENSG00000144306	SCRN3	secernin 3 [Source:HGNC Symbol;Acc:HGNC:30382]	-0.86566	0.004199
ENSG00000100983	GSS	glutathione synthetase [Source:HGNC Symbol;Acc:HGNC:4624]	-0.85585	0.004204
ENSG00000182870	GALNT9	polypeptide N-acetylgalactosaminyltransferase 9 [Source:HGNC Symbol;Acc:HGNC:4131]	-0.85098	0.004213
ENSG00000256159			-1.05229	0.004217
ENSG00000169914	OTUD3	OTU deubiquitinase 3 [Source:HGNC Symbol;Acc:HGNC:29038]	-1.0518	0.004225
ENSG00000107807	TLX1	T-cell leukemia homeobox 1 [Source:HGNC Symbol;Acc:HGNC:5056]	-1.04966	0.004265
ENSG00000230487	PSMG3-AS1	PSMG3 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:22230]	-1.04955	0.004274
ENSG00000107798	LIPA	lipase A, lysosomal acid type [Source:HGNC Symbol;Acc:HGNC:6617]	-0.85548	0.004294
ENSG00000226264	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000234154	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000239329	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000241296	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000241674	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000242092	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000242386	HLA-DMB	major histocompatibility complex, class II, DM beta [Source:HGNC Symbol;Acc:HGNC:4935]	-0.94579	0.004313
ENSG00000251595	ABCA11P	ATP binding cassette subfamily A member 11, pseudogene [Source:HGNC Symbol;Acc:HGNC:31]	-0.94706	0.004323

ENSG00000143630	HCN3	hyperpolarization activated cyclic nucleotide gated potassium channel 3 [Source:HGNC Symbol;Acc:HGNC:19183]	-1.08467	0.004338
ENSG00000050767	COL23A1	collagen type XXIII alpha 1 chain [Source:HGNC Symbol;Acc:HGNC:22990]	-1.04696	0.004354
ENSG00000179447			-1.04689	0.004357
ENSG00000082397	EPB41L3	erythrocyte membrane protein band 4.1 like 3 [Source:HGNC Symbol;Acc:HGNC:3380]	-0.98494	0.00437
ENSG00000146373	RNF217	ring finger protein 217 [Source:HGNC Symbol;Acc:HGNC:21487]	-0.89238	0.004398
ENSG00000055483	USP36	ubiquitin specific peptidase 36 [Source:HGNC Symbol;Acc:HGNC:20062]	-1.04616	0.004423
ENSG00000197170	PSMD12	proteasome 26S subunit, non-ATPase 12 [Source:HGNC Symbol;Acc:HGNC:9557]	-0.84644	0.004501
ENSG00000176868			-1.04084	0.004576
ENSG00000227354	RBM26-AS1	RBM26 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:39805]	-1.04066	0.004612
ENSG00000170162	VGLL2	vestigial like family member 2 [Source:HGNC Symbol;Acc:HGNC:20232]	-0.87792	0.004613
ENSG00000145476	CYP4V2	cytochrome P450 family 4 subfamily V member 2 [Source:HGNC Symbol;Acc:HGNC:23198]	-0.9355	0.004616
ENSG00000182841	RRP7BP	ribosomal RNA processing 7 homolog B, pseudogene [Source:HGNC Symbol;Acc:HGNC:30454]	-0.93776	0.004617
ENSG00000213963			-1.03918	0.004624
ENSG00000116039	ATP6V1B1	ATPase H+ transporting V1 subunit B1 [Source:HGNC Symbol;Acc:HGNC:853]	-0.83944	0.004633
ENSG00000167608	TMC4	transmembrane channel like 4 [Source:HGNC Symbol;Acc:HGNC:22998]	-0.8449	0.004648
ENSG00000141569	TRIM65	tripartite motif containing 65 [Source:HGNC Symbol;Acc:HGNC:27316]	-1.03891	0.004654
ENSG00000171295	ZNF440	zinc finger protein 440 [Source:HGNC Symbol;Acc:HGNC:20874]	-1.03947	0.004654
ENSG00000217085	HMGB3P19	high mobility group box 3 pseudogene 19 [Source:HGNC Symbol;Acc:HGNC:39311]	-1.03769	0.004687
ENSG00000156170	NDUFAF6	NADH:ubiquinone oxidoreductase complex assembly factor 6 [Source:HGNC Symbol;Acc:HGNC:28625]	-0.97478	0.00471
ENSG00000105948	TTC26	tetratricopeptide repeat domain 26 [Source:HGNC Symbol;Acc:HGNC:21882]	-1.10835	0.004729
ENSG00000239911	PRKAG2-AS1	PRKAG2 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:40468]	-1.03578	0.004741
ENSG00000259116			-1.03579	0.004741
ENSG00000269947			-1.03594	0.004745
ENSG00000198482	ZNF808	zinc finger protein 808 [Source:HGNC Symbol;Acc:HGNC:33230]	-0.87446	0.004752
ENSG00000225077	LINC00337	long intergenic non-protein coding RNA 337 [Source:HGNC Symbol;Acc:HGNC:28620]	-1.03519	0.004767
ENSG00000182578	CSF1R	colony stimulating factor 1 receptor [Source:HGNC Symbol;Acc:HGNC:2433]	-0.86402	0.004787
ENSG00000185513	L3MBTL1	l(3)mbt-like 1 (Drosophila) [Source:HGNC Symbol;Acc:HGNC:15905]	-0.93054	0.004844
ENSG00000267150			-1.03248	0.004861
ENSG00000070882	OSBPL3	oxysterol binding protein like 3 [Source:HGNC Symbol;Acc:HGNC:16370]	-1.08649	0.004879
ENSG00000239665			-1.03228	0.00488
ENSG00000257097	CLIP1-AS1	CLIP1 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:48586]	-1.03078	0.00493
ENSG00000229509			-2.65872	0.004937
ENSG00000270767			-2.65872	0.004937
ENSG00000271644			-2.65872	0.004937
ENSG00000206072	SERPINB11	serpin family B member 11 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:14221]	-0.83385	0.004951
ENSG00000275517	PRB1	proline rich protein BstNI subfamily 1 [Source:HGNC Symbol;Acc:HGNC:9337]	-0.94022	0.005055

ENSG00000154016	GRAP	GRB2-related adaptor protein [Source:HGNC Symbol;Acc:HGNC:4562]	-0.83476	0.005082
ENSG00000253181			-1.02618	0.0051
ENSG00000214226	C17orf67	chromosome 17 open reading frame 67 [Source:HGNC Symbol;Acc:HGNC:27900]	-1.02641	0.005122
ENSG00000268858			-1.02734	0.00514
ENSG00000092068	SLC7A8	solute carrier family 7 member 8 [Source:HGNC Symbol;Acc:HGNC:11066]	-0.83031	0.00515
ENSG00000253741			-1.02422	0.005179
ENSG00000137434	C6orf52	chromosome 6 open reading frame 52 [Source:HGNC Symbol;Acc:HGNC:20881]	-1.02357	0.005201
ENSG00000135679	MDM2	MDM2 proto-oncogene [Source:HGNC Symbol;Acc:HGNC:6973]	-1.01222	0.005225
ENSG00000205129	C4orf47	chromosome 4 open reading frame 47 [Source:HGNC Symbol;Acc:HGNC:34346]	-1.02318	0.005228
ENSG00000231365			-1.02286	0.00523
ENSG00000206454	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]	-0.93115	0.005266
ENSG00000229094	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]	-0.93115	0.005266
ENSG00000230336	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]	-0.93115	0.005266
ENSG00000233911	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]	-0.93115	0.005266
ENSG00000235068	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]	-0.93115	0.005266
ENSG00000237582	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]	-0.93115	0.005266
ENSG00000258881			-1.02072	0.005315
ENSG00000251474	RPL32P3	ribosomal protein L32 pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:27024]	-0.92182	0.005342
ENSG00000253669			-1.02008	0.005373
ENSG00000166136	NDUFB8	NADH:ubiquinone oxidoreductase subunit B8 [Source:HGNC Symbol;Acc:HGNC:7703]	-0.92175	0.005384
ENSG00000205181	LINC00654	long intergenic non-protein coding RNA 654 [Source:HGNC Symbol;Acc:HGNC:27154]	-1.01856	0.005409
ENSG00000197121	PGAP1	post-GPI attachment to proteins 1 [Source:HGNC Symbol;Acc:HGNC:25712]	-0.82561	0.005417
ENSG00000244457	ENO1P1	enolase 1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:3352]	-1.02796	0.005471
ENSG00000010327	STAB1	stabilin 1 [Source:HGNC Symbol;Acc:HGNC:18628]	-0.82518	0.005477
ENSG00000170340	B3GNT2	UDP-GlcNAc:betaGal beta-1,3-N- acetylglucosaminyltransferase 2 [Source:HGNC Symbol;Acc:HGNC:15629]	-0.82441	0.005558
ENSG00000128250	RFPL1	ret finger protein like 1 [Source:HGNC Symbol;Acc:HGNC:9977]	-1.01411	0.005587
ENSG00000182722	SEPHS1P1	selenophosphate synthetase 1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:42161]	-0.91375	0.005623
ENSG00000148942	SLC5A12	solute carrier family 5 member 12 [Source:HGNC Symbol;Acc:HGNC:28750]	-0.81979	0.005651
ENSG00000124486	USP9X	ubiquitin specific peptidase 9, X-linked [Source:HGNC Symbol;Acc:HGNC:12632]	-0.87039	0.005655
ENSG00000072364	AFF4	AF4/FMR2 family member 4 [Source:HGNC Symbol;Acc:HGNC:17869]	-0.87746	0.005682
ENSG00000167194	C16orf92	chromosome 16 open reading frame 92 [Source:HGNC Symbol;Acc:HGNC:26346]	-1.01182	0.005685
ENSG00000185619	PCGF3	polycomb group ring finger 3 [Source:HGNC Symbol;Acc:HGNC:10066]	-1.01457	0.005697
ENSG00000249230	CDH12P2	cadherin 12 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:1752]	-2.5935	0.005705
ENSG00000280638	CDH12P2	cadherin 12 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:1752]	-2.5935	0.005705
ENSG00000105341	ATP5SL	ATP5S like [Source:HGNC Symbol;Acc:HGNC:25496]	-0.82216	0.00571
ENSG00000214273	AGGF1P1	angiogenic factor with G-patch and FHA domains 1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:38051]	-1.01133	0.005716

ENSG00000233435	AGGF1P2	angiogenic factor with G-patch and FHA domains 1 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:38052]	-1.01133	0.005716
ENSG00000150667	FSIP1	fibrous sheath interacting protein 1 [Source:HGNC Symbol;Acc:HGNC:21674]	-1.25459	0.005751
ENSG00000273676			-1.00896	0.005808
ENSG00000274516	FAM74A6	family with sequence similarity 74 member A6 [Source:HGNC Symbol;Acc:HGNC:34036]	-1.00896	0.005808
ENSG00000274583	FAM74A4	family with sequence similarity 74 member A4 [Source:HGNC Symbol;Acc:HGNC:32032]	-1.00896	0.005808
ENSG00000229972	IQCF3	IQ motif containing F3 [Source:HGNC Symbol;Acc:HGNC:31816]	-1.00852	0.005829
ENSG00000099251	HSD17B7P2	hydroxysteroid 17-beta dehydrogenase 7 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:28120]	-0.91096	0.005849
ENSG00000160963	COL26A1	collagen type XXVI alpha 1 chain [Source:HGNC Symbol;Acc:HGNC:18038]	-1.00815	0.005851
ENSG00000163975	MELTF	melanotransferrin [Source:HGNC Symbol;Acc:HGNC:7037]	-0.81654	0.005868
ENSG00000188199	NUTM2B	NUT family member 2B [Source:HGNC Symbol;Acc:HGNC:23445]	-1.0075	0.005875
ENSG00000228570	NUTM2E	NUT family member 2E [Source:HGNC Symbol;Acc:HGNC:23448]	-1.0075	0.005875
ENSG00000263069			-1.0072	0.005886
ENSG00000197837	HIST4H4	histone cluster 4 H4 [Source:HGNC Symbol;Acc:HGNC:20510]	-0.81419	0.005914
ENSG00000168405	CMAHP	cytidine monophospho-N-acetylneuraminic acid hydroxylase, pseudogene [Source:HGNC Symbol;Acc:HGNC:2098]	-1.00661	0.005915
ENSG00000274254	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000274394	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000274480	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000274511	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000274696	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000274763	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000274786	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000275513	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000276084	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000276196	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000276328	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000276433	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000276572	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932

ENSG00000277552	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000277596	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000277620	KIR3DL3	killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 3 [Source:HGNC Symbol;Acc:HGNC:16312]	-1.00618	0.005932
ENSG00000259158	ADAM20P1	ADAM metalloproteinase domain 20 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:20102]	-1.00613	0.005933
ENSG00000254076			-1.00572	0.005953
ENSG00000224356			-1.0054	0.005965
ENSG00000148187	MRRF	mitochondrial ribosome recycling factor [Source:HGNC Symbol;Acc:HGNC:7234]	-1.08998	0.005975
ENSG00000183458			-1.00642	0.006032
ENSG00000244257	PKD1P1	polycystin 1, transient receptor potential channel interacting pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:30065]	-1.00642	0.006032
ENSG00000235827	TUBB8P9	tubulin beta 8 class VIII pseudogene 9 [Source:HGNC Symbol;Acc:HGNC:42347]	-1.00593	0.006047
ENSG00000237469	TUBB8P10	tubulin beta 8 class VIII pseudogene 10 [Source:HGNC Symbol;Acc:HGNC:42348]	-1.00593	0.006047
ENSG00000162104	ADCY9	adenylate cyclase 9 [Source:HGNC Symbol;Acc:HGNC:240]	-0.81668	0.006048
ENSG00000135094	SDS	serine dehydratase [Source:HGNC Symbol;Acc:HGNC:10691]	-0.81617	0.006118
ENSG00000196814	MVB12B	multivesicular body subunit 12B [Source:HGNC Symbol;Acc:HGNC:23368]	-0.80983	0.006122
ENSG00000196782	MAML3	mastermind like transcriptional coactivator 3 [Source:HGNC Symbol;Acc:HGNC:16272]	-1.00129	0.006156
ENSG00000125462	C1orf61	chromosome 1 open reading frame 61 [Source:HGNC Symbol;Acc:HGNC:30780]	-0.81149	0.00627
ENSG00000211807	TRAV26-1	T-cell receptor alpha variable 26-1 [Source:HGNC Symbol;Acc:HGNC:12123]	-0.99817	0.006298
ENSG00000129467	ADCY4	adenylate cyclase 4 [Source:HGNC Symbol;Acc:HGNC:235]	-0.80927	0.006369
ENSG00000099282	TSPAN15	tetraspanin 15 [Source:HGNC Symbol;Acc:HGNC:23298]	-0.80778	0.006416
ENSG00000214711	CAPN14	calpain 14 [Source:HGNC Symbol;Acc:HGNC:16664]	-0.99513	0.006443
ENSG00000177575	CD163	CD163 molecule [Source:HGNC Symbol;Acc:HGNC:1631]	-0.80647	0.006467
ENSG00000007129	CEACAM21	carcinoembryonic antigen related cell adhesion molecule 21 [Source:HGNC Symbol;Acc:HGNC:28834]	-0.89835	0.006469
ENSG00000278565	CEACAM21	carcinoembryonic antigen related cell adhesion molecule 21 [Source:HGNC Symbol;Acc:HGNC:28834]	-0.89835	0.006469
ENSG00000130413	STK33	serine/threonine kinase 33 [Source:HGNC Symbol;Acc:HGNC:14568]	-0.80855	0.006474
ENSG00000231738	TSPAN19	tetraspanin 19 [Source:HGNC Symbol;Acc:HGNC:31886]	-0.99422	0.006488
ENSG00000171405	XAGE5	X antigen family member 5 [Source:HGNC Symbol;Acc:HGNC:30930]	-0.80426	0.006571
ENSG00000124602	UNC5CL	unc-5 family C-terminal like [Source:HGNC Symbol;Acc:HGNC:21203]	-0.90902	0.00659
ENSG00000079102	RUNX1T1	RUNX1 translocation partner 1 [Source:HGNC Symbol;Acc:HGNC:1535]	-0.79952	0.006624
ENSG00000141040	ZNF287	zinc finger protein 287 [Source:HGNC Symbol;Acc:HGNC:13502]	-0.80834	0.006676
ENSG00000102393	GLA	galactosidase alpha [Source:HGNC Symbol;Acc:HGNC:4296]	-0.80141	0.006679
ENSG00000173110	HSPA6	heat shock protein family A (Hsp70) member 6 [Source:HGNC Symbol;Acc:HGNC:5239]	-0.99016	0.006686
ENSG00000248810			-0.98855	0.006767

ENSG0000011566	MAP4K3	mitogen-activated protein kinase kinase kinase kinase 3 [Source:HGNC Symbol;Acc:HGNC:6865]	-0.79911	0.006792
ENSG00000225684	FAM225B	family with sequence similarity 225 member B (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:21865]	-0.98792	0.006832
ENSG00000142168	SOD1	superoxide dismutase 1 [Source:HGNC Symbol;Acc:HGNC:11179]	0.804374	0.006863
ENSG00000140632	GLYR1	glyoxylate reductase 1 homolog [Source:HGNC Symbol;Acc:HGNC:24434]	-0.98618	0.006911
ENSG00000146192	FGD2	FYVE, RhoGEF and PH domain containing 2 [Source:HGNC Symbol;Acc:HGNC:3664]	-0.79922	0.006934
ENSG00000275555	DLL1	delta like canonical Notch ligand 1 [Source:HGNC Symbol;Acc:HGNC:2908]	-2.50214	0.007002
ENSG00000235169	SMIM1	small integral membrane protein 1 (Vel blood group) [Source:HGNC Symbol;Acc:HGNC:44204]	-0.98383	0.007016
ENSG00000256762	STH	saitohin [Source:HGNC Symbol;Acc:HGNC:18839]	-0.8873	0.007053
ENSG00000281139	STH	saitohin [Source:HGNC Symbol;Acc:HGNC:18839]	-0.8873	0.007053
ENSG00000201659	RNU12-2P	RNA, U12 small nuclear 2, pseudogene [Source:HGNC Symbol;Acc:HGNC:10109]	-0.98334	0.007059
ENSG00000112299	VNN1	vanin 1 [Source:HGNC Symbol;Acc:HGNC:12705]	-0.8634	0.007062
ENSG00000170469	SPATA24	spermatogenesis associated 24 [Source:HGNC Symbol;Acc:HGNC:27322]	-0.98524	0.007073
ENSG00000165124	SVEP1	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1 [Source:HGNC Symbol;Acc:HGNC:15985]	-0.80068	0.007097
ENSG00000227300	KRT16P2	keratin 16 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:37807]	-0.98229	0.00711
ENSG00000156414	TDRD9	tudor domain containing 9 [Source:HGNC Symbol;Acc:HGNC:20122]	-0.96693	0.007121
ENSG00000242294	STAG3L5P	stromal antigen 3-like 5 pseudogene [Source:HGNC Symbol;Acc:HGNC:48896]	-2.49454	0.007123
ENSG00000114631	PODXL2	podocalyxin like 2 [Source:HGNC Symbol;Acc:HGNC:17936]	-0.79649	0.007138
ENSG00000105376	ICAM5	intercellular adhesion molecule 5 [Source:HGNC Symbol;Acc:HGNC:5348]	-0.79416	0.007233
ENSG00000229927	RHEBP1	Ras-homolog enriched in brain pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:10010]	-2.48351	0.007302
ENSG00000172493	AFF1	AF4/FMR2 family member 1 [Source:HGNC Symbol;Acc:HGNC:7135]	-0.80767	0.007346
ENSG00000282880	B3GALNT2	beta-1,3-N-acetylgalactosaminyltransferase 2 [Source:HGNC Symbol;Acc:HGNC:28596]	-0.88442	0.007355
ENSG00000206467	OR2H2	olfactory receptor family 2 subfamily H member 2 [Source:HGNC Symbol;Acc:HGNC:8253]	-0.88211	0.007444
ENSG00000224319	OR2H2	olfactory receptor family 2 subfamily H member 2 [Source:HGNC Symbol;Acc:HGNC:8253]	-0.88211	0.007444
ENSG00000227044	OR2H2	olfactory receptor family 2 subfamily H member 2 [Source:HGNC Symbol;Acc:HGNC:8253]	-0.88211	0.007444
ENSG00000229185	OR2H2	olfactory receptor family 2 subfamily H member 2 [Source:HGNC Symbol;Acc:HGNC:8253]	-0.88211	0.007444
ENSG00000146469	VIP	vasoactive intestinal peptide [Source:HGNC Symbol;Acc:HGNC:12693]	-0.78951	0.007461
ENSG00000088179	PTPN4	protein tyrosine phosphatase, non-receptor type 4 [Source:HGNC Symbol;Acc:HGNC:9656]	-1.22171	0.007466
ENSG00000130513	GDF15	growth differentiation factor 15 [Source:HGNC Symbol;Acc:HGNC:30142]	-1.11918	0.007483
ENSG00000282608	ADORA3	adenosine A3 receptor [Source:HGNC Symbol;Acc:HGNC:268]	-0.882	0.007529
ENSG00000248586			-0.97353	0.007595
ENSG00000164331	ANKRA2	ankyrin repeat family A member 2 [Source:HGNC Symbol;Acc:HGNC:13208]	-0.78828	0.007604
ENSG00000282788			-0.88911	0.007606

ENSG00000095970	TREM2	triggering receptor expressed on myeloid cells 2 [Source:HGNC Symbol;Acc:HGNC:17761]	-0.78919	0.007621
ENSG00000100796	PPP4R3A	protein phosphatase 4 regulatory subunit 3A [Source:HGNC Symbol;Acc:HGNC:20219]	-0.79234	0.007645
ENSG00000279851			-2.46229	0.007662
ENSG00000155034	FBXL18	F-box and leucine rich repeat protein 18 [Source:HGNC Symbol;Acc:HGNC:21874]	-0.94824	0.00769
ENSG00000253628			-0.97169	0.007705
ENSG00000174013	FBXO45	F-box protein 45 [Source:HGNC Symbol;Acc:HGNC:29148]	-0.97157	0.007707
ENSG00000235978			-0.97105	0.007712
ENSG00000203711	C6orf99	chromosome 6 open reading frame 99 [Source:HGNC Symbol;Acc:HGNC:21179]	-0.97098	0.007713
ENSG00000131738	KRT33B	keratin 33B [Source:HGNC Symbol;Acc:HGNC:6451]	-0.79161	0.007745
ENSG00000255559	ZNF252P-AS1	ZNF252P antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:27821]	-0.97064	0.007764
ENSG00000081154	PCNP	PEST proteolytic signal containing nuclear protein [Source:HGNC Symbol;Acc:HGNC:30023]	-0.78895	0.007766
ENSG00000235527	HIPK1-AS1	HIPK1 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:50576]	-0.97004	0.007768
ENSG00000197549	PRAMENP	PRAME N-terminal-like, pseudogene [Source:HGNC Symbol;Acc:HGNC:34302]	-0.96989	0.007773
ENSG00000205786	LINC01531	long intergenic non-protein coding RNA 1531 [Source:HGNC Symbol;Acc:HGNC:51270]	-0.9696	0.007792
ENSG00000108813	DLX4	distal-less homeobox 4 [Source:HGNC Symbol;Acc:HGNC:2917]	-0.78887	0.007793
ENSG00000277222	BTBD7	BTB domain containing 7 [Source:HGNC Symbol;Acc:HGNC:18269]	-0.87891	0.007793
ENSG00000273884	LILRA3	leukocyte immunoglobulin like receptor A3 [Source:HGNC Symbol;Acc:HGNC:6604]	-1.37477	0.00784
ENSG00000275841	LILRA3	leukocyte immunoglobulin like receptor A3 [Source:HGNC Symbol;Acc:HGNC:6604]	-1.37477	0.00784
ENSG00000276175	LILRA3	leukocyte immunoglobulin like receptor A3 [Source:HGNC Symbol;Acc:HGNC:6604]	-1.37477	0.00784
ENSG00000278046	LILRA3	leukocyte immunoglobulin like receptor A3 [Source:HGNC Symbol;Acc:HGNC:6604]	-1.37477	0.00784
ENSG00000155008	APOOL	apolipoprotein O like [Source:HGNC Symbol;Acc:HGNC:24009]	-0.9662	0.007861
ENSG00000206487	C6orf136	chromosome 6 open reading frame 136 [Source:HGNC Symbol;Acc:HGNC:21301]	-0.88036	0.007872
ENSG00000224120	C6orf136	chromosome 6 open reading frame 136 [Source:HGNC Symbol;Acc:HGNC:21301]	-0.88036	0.007872
ENSG00000233164	C6orf136	chromosome 6 open reading frame 136 [Source:HGNC Symbol;Acc:HGNC:21301]	-0.88036	0.007872
ENSG00000233641	C6orf136	chromosome 6 open reading frame 136 [Source:HGNC Symbol;Acc:HGNC:21301]	-0.88036	0.007872
ENSG00000237012	C6orf136	chromosome 6 open reading frame 136 [Source:HGNC Symbol;Acc:HGNC:21301]	-0.88036	0.007872
ENSG00000237100	C6orf136	chromosome 6 open reading frame 136 [Source:HGNC Symbol;Acc:HGNC:21301]	-0.88036	0.007872
ENSG00000221900	POM121L12	POM121 transmembrane nucleoporin like 12 [Source:HGNC Symbol;Acc:HGNC:25369]	-0.78587	0.007896
ENSG00000224934			-0.96754	0.007909
ENSG00000224075	TTY2	testis-specific transcript, Y-linked 22 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:18846]	-0.96803	0.007934
ENSG00000243479	MNX1-AS1	MNX1 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48954]	-0.96715	0.007939
ENSG00000276262	NSF	N-ethylmaleimide sensitive factor, vesicle fusing ATPase [Source:HGNC Symbol;Acc:HGNC:8016]	-0.87342	0.00797
ENSG00000278174	NSF	N-ethylmaleimide sensitive factor, vesicle fusing ATPase [Source:HGNC Symbol;Acc:HGNC:8016]	-0.87342	0.00797

ENSG00000263296	KRTAP3-2	keratin associated protein 3-2 [Source:HGNC Symbol;Acc:HGNC:16779]	-0.87229	0.008009
ENSG00000272812			-2.44249	0.008014
ENSG00000233839			-0.96615	0.008019
ENSG00000104497	SNX16	sorting nexin 16 [Source:HGNC Symbol;Acc:HGNC:14980]	-0.78362	0.008021
ENSG00000181754	AMIGO1	adhesion molecule with Ig like domain 1 [Source:HGNC Symbol;Acc:HGNC:20824]	-0.78418	0.008023
ENSG00000160886	LY6K	lymphocyte antigen 6 family member K [Source:HGNC Symbol;Acc:HGNC:24225]	-0.80748	0.008055
ENSG00000163428	LRR58	leucine rich repeat containing 58 [Source:HGNC Symbol;Acc:HGNC:26968]	-0.96513	0.00807
ENSG00000240068	RPL21P42	ribosomal protein L21 pseudogene 42 [Source:HGNC Symbol;Acc:HGNC:37017]	-0.96426	0.008107
ENSG00000240138	EEF1G4	eukaryotic translation elongation factor 1 gamma pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:44559]	-0.96419	0.008109
ENSG00000279386			-0.96337	0.008156
ENSG00000187730	GABRD	gamma-aminobutyric acid type A receptor delta subunit [Source:HGNC Symbol;Acc:HGNC:4084]	-0.79797	0.008186
ENSG00000205746			-0.96322	0.008187
ENSG00000227827			-0.96322	0.008187
ENSG00000254681	PKD1P5	polycystin 1, transient receptor potential channel interacting pseudogene 5 [Source:HGNC Symbol;Acc:HGNC:30069]	-0.96322	0.008187
ENSG00000282522	PKD1P1	polycystin 1, transient receptor potential channel interacting pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:30065]	-0.96322	0.008187
ENSG00000197429	IPP	intracisternal A particle-promoted polypeptide [Source:HGNC Symbol;Acc:HGNC:6108]	-0.96251	0.008218
ENSG00000233471	KRT18P62	keratin 18 pseudogene 62 [Source:HGNC Symbol;Acc:HGNC:48889]	-0.96383	0.008221
ENSG00000137460	FHDC1	FH2 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29363]	-0.88293	0.008238
ENSG00000244226	ILF2P1	interleukin enhancer binding factor 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:44695]	-0.96272	0.008265
ENSG00000272145	NFYC-AS1	NFYC antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:49451]	-0.96015	0.008351
ENSG00000183250	LINC01547	long intergenic non-protein coding RNA 1547 [Source:HGNC Symbol;Acc:HGNC:15707]	-0.7799	0.00837
ENSG00000079101	CLUL1	clusterin like 1 [Source:HGNC Symbol;Acc:HGNC:2096]	-0.85764	0.0084
ENSG00000103226	NOMO3	NODAL modulator 3 [Source:HGNC Symbol;Acc:HGNC:25242]	-2.42072	0.008421
ENSG00000278087	NOMO3	NODAL modulator 3 [Source:HGNC Symbol;Acc:HGNC:25242]	-2.42072	0.008421
ENSG00000124610	HIST1H1A	histone cluster 1 H1 family member a [Source:HGNC Symbol;Acc:HGNC:4715]	-0.78335	0.008444
ENSG00000254502			-0.95975	0.008446
ENSG00000280047			-0.95877	0.008474
ENSG00000215174	NLRP2B	NLR family pyrin domain containing 2B [Source:HGNC Symbol;Acc:HGNC:29887]	-0.95812	0.008481
ENSG00000264522	OTUD7B	OTU deubiquitinase 7B [Source:HGNC Symbol;Acc:HGNC:16683]	-0.95807	0.008498
ENSG00000213483	NDUF4P2	NADH:ubiquinone oxidoreductase complex assembly factor 4 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:44541]	-0.95758	0.008527
ENSG00000134339	SAA2	serum amyloid A2 [Source:HGNC Symbol;Acc:HGNC:10514]	-2.41488	0.008534
ENSG00000152457	DCLRE1C	DNA cross-link repair 1C [Source:HGNC Symbol;Acc:HGNC:17642]	-0.77605	0.008534
ENSG00000267059			-0.95712	0.00854
ENSG00000129455	KLK8	kallikrein related peptidase 8 [Source:HGNC Symbol;Acc:HGNC:6369]	-0.87226	0.008544

ENSG00000279331	RBM12B-AS1	RBM12B antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:28818]	-0.95656	0.008598
ENSG00000206530	CFAP44	cilia and flagella associated protein 44 [Source:HGNC Symbol;Acc:HGNC:25631]	-1.01051	0.008698
ENSG00000149262	INTS4	integrator complex subunit 4 [Source:HGNC Symbol;Acc:HGNC:25048]	-0.77457	0.00877
ENSG00000206418	RAB12	RAB12, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:31332]	-0.86885	0.00879
ENSG00000224546	EIF4BP3	eukaryotic translation initiation factor 4B pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:37936]	-0.95988	0.008793
ENSG00000180777	ANKRD30B	ankyrin repeat domain 30B [Source:HGNC Symbol;Acc:HGNC:24165]	-2.39949	0.008839
ENSG00000138399	FASTKD1	FAST kinase domains 1 [Source:HGNC Symbol;Acc:HGNC:26150]	-0.78128	0.008867
ENSG00000146001	PCDHB18P	protocadherin beta 18 pseudogene [Source:HGNC Symbol;Acc:HGNC:14548]	-0.86048	0.008892
ENSG00000173557	C2orf70	chromosome 2 open reading frame 70 [Source:HGNC Symbol;Acc:HGNC:27938]	-0.95136	0.008908
ENSG00000161640	SIGLEC11	sialic acid binding Ig like lectin 11 [Source:HGNC Symbol;Acc:HGNC:15622]	-0.90253	0.008922
ENSG00000213088	ACKR1	atypical chemokine receptor 1 (Duffy blood group) [Source:HGNC Symbol;Acc:HGNC:4035]	-0.76981	0.008925
ENSG00000113494	PRLR	prolactin receptor [Source:HGNC Symbol;Acc:HGNC:9446]	-0.7729	0.008943
ENSG00000124444	ZNF576	zinc finger protein 576 [Source:HGNC Symbol;Acc:HGNC:28357]	-0.86362	0.009017
ENSG00000282992	LGALS4	galectin 4 [Source:HGNC Symbol;Acc:HGNC:6565]	-0.86752	0.009024
ENSG00000142396	ERVK3-1	endogenous retrovirus group K3 member 1 [Source:HGNC Symbol;Acc:HGNC:30466]	-2.39016	0.00903
ENSG00000185182	GOLGA8DP	golgin A8 family member D, pseudogene [Source:HGNC Symbol;Acc:HGNC:32376]	-1.11349	0.009038
ENSG00000261554			-0.95024	0.009042
ENSG00000138686	BBS7	Bardet-Biedl syndrome 7 [Source:HGNC Symbol;Acc:HGNC:18758]	-0.79113	0.009051
ENSG00000229522	LINC01523	long intergenic non-protein coding RNA 1523 [Source:HGNC Symbol;Acc:HGNC:51225]	-0.94886	0.009074
ENSG00000139679	LPAR6	lysophosphatidic acid receptor 6 [Source:HGNC Symbol;Acc:HGNC:15520]	-0.77517	0.009098
ENSG00000173705	SUSD5	sushi domain containing 5 [Source:HGNC Symbol;Acc:HGNC:29061]	-0.9495	0.009113
ENSG00000134574	DDB2	damage specific DNA binding protein 2 [Source:HGNC Symbol;Acc:HGNC:2718]	-0.89953	0.009134
ENSG00000204650	LINC02210	long intergenic non-protein coding RNA 2210 [Source:HGNC Symbol;Acc:HGNC:26327]	-0.85562	0.00923
ENSG00000282171		CRHR1 intronic transcript 1 [Source:EntrezGene;Acc:147081]	-0.85562	0.00923
ENSG00000155659	VSIG4	V-set and immunoglobulin domain containing 4 [Source:HGNC Symbol;Acc:HGNC:17032]	-0.7706	0.009316
ENSG00000273707	CDKN1C	cyclin dependent kinase inhibitor 1C [Source:HGNC Symbol;Acc:HGNC:1786]	2.37637	0.009319
ENSG00000188038	NRN1L	neurtin 1 like [Source:HGNC Symbol;Acc:HGNC:29811]	-0.85519	0.009337
ENSG00000188032	C19orf67	chromosome 19 open reading frame 67 [Source:HGNC Symbol;Acc:HGNC:34354]	-0.9448	0.009356
ENSG00000166979	EVA1C	eva-1 homolog C [Source:HGNC Symbol;Acc:HGNC:13239]	-0.87452	0.009395
ENSG00000196922	ZNF252P	zinc finger protein 252, pseudogene [Source:HGNC Symbol;Acc:HGNC:13046]	-0.94516	0.009407
ENSG00000267296	CEBPA-AS1	CEBPA antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:25710]	-0.94386	0.009422
ENSG00000264845			-0.94329	0.009516
ENSG00000215325	ASS1P10	argininosuccinate synthetase 1 pseudogene 10 [Source:HGNC Symbol;Acc:HGNC:760]	-0.94227	0.009523
ENSG00000274113	LILRA5	leukocyte immunoglobulin like receptor A5 [Source:HGNC Symbol;Acc:HGNC:16309]	-0.85697	0.009568

ENSG00000274914	LILRA5	leukocyte immunoglobulin like receptor A5 [Source:HGNC Symbol;Acc:HGNC:16309]	-0.85697	0.009568
ENSG00000275404	LILRA5	leukocyte immunoglobulin like receptor A5 [Source:HGNC Symbol;Acc:HGNC:16309]	-0.85697	0.009568
ENSG00000278355	LILRA5	leukocyte immunoglobulin like receptor A5 [Source:HGNC Symbol;Acc:HGNC:16309]	-0.85697	0.009568
ENSG00000254469		Putative short transient receptor potential channel 2-like protein [Source:UniProtKB/Swiss- Prot;Acc:Q6ZNB5]	-0.94156	0.009583
ENSG00000086712	TXLNG	taxilin gamma [Source:HGNC Symbol;Acc:HGNC:18578]	-0.94214	0.009591
ENSG00000147257	GPC3	glypican 3 [Source:HGNC Symbol;Acc:HGNC:4451]	-0.76544	0.009612
ENSG00000245849	RAD51-AS1	RAD51 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48621]	-0.94232	0.009613
ENSG00000113569	NUP155	nucleoporin 155 [Source:HGNC Symbol;Acc:HGNC:8063]	-0.76254	0.009622
ENSG00000214776			-0.94152	0.009625
ENSG00000203907	OOEP	oocyte expressed protein [Source:HGNC Symbol;Acc:HGNC:21382]	-0.94056	0.009652
ENSG00000251510			-0.94037	0.009668
ENSG00000230562	FAM133DP	family with sequence similarity 133 member D, pseudogene [Source:HGNC Symbol;Acc:HGNC:44192]	-0.93955	0.009711
ENSG00000248485	PCP4L1	Purkinje cell protein 4 like 1 [Source:HGNC Symbol;Acc:HGNC:20448]	-0.94002	0.009721
ENSG00000172318	B3GALT1	beta-1,3-galactosyltransferase 1 [Source:HGNC Symbol;Acc:HGNC:916]	-0.76423	0.009735
ENSG00000232160	RAP2C-AS1	RAP2C antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:40957]	-0.93996	0.009759
ENSG00000169116	PARM1	prostate androgen-regulated mucin-like protein 1 [Source:HGNC Symbol;Acc:HGNC:24536]	-0.86631	0.009822
ENSG00000211714	TRBV7-3	T-cell receptor beta variable 7-3 [Source:HGNC Symbol;Acc:HGNC:12237]	-0.93847	0.009835
ENSG00000135426	TESPA1	thymocyte expressed, positive selection associated 1 [Source:HGNC Symbol;Acc:HGNC:29109]	-0.93731	0.009877
ENSG00000138434	SSFA2	sperm specific antigen 2 [Source:HGNC Symbol;Acc:HGNC:11319]	-0.7627	0.009883
ENSG00000257528	KRT8P19	keratin 8 pseudogene 19 [Source:HGNC Symbol;Acc:HGNC:33371]	-0.93775	0.00989
ENSG00000121940	CLCC1	chloride channel CLIC like 1 [Source:HGNC Symbol;Acc:HGNC:29675]	-0.84667	0.009893
ENSG00000261706	LINC00165	long intergenic non-protein coding RNA 165 [Source:HGNC Symbol;Acc:HGNC:33166]	-0.93699	0.009949
ENSG00000278463	HIST1H2AB	histone cluster 1 H2A family member b [Source:HGNC Symbol;Acc:HGNC:4734]	-0.75817	0.009977

Πίνακας 9: Γονίδια που ξεχώρισαν ως διαφορετικά εκφρασμένα μέσω μετα-ανάλυσης, με βάση τιμή $p\text{-value} < 0.01$, έπειτα από έκθεση σε χαμηλή δόση ακτινοβολίας ($< 0.5\text{Gy}$) στο διάστημα 4-16 ωρών. Συμπεριλαμβάνεται το αντίστοιχο HGNC σύμβολο με την αντίστοιχη περιγραφή και οι ακριβείς τιμές του υπολογισμένου $p\text{-value}$ και της διαφοράς στην έκφραση.

ENSG	HGNC	Description	Diff	P-value
ENSG00000268858			-2.15717	5.20E-07
ENSG00000204437	CTSLP6	cathepsin L pseudogene 6 [Source:HGNC Symbol;Acc:HGNC:23611]	-1.73566	0.000015
ENSG00000230011	CTSLP4	cathepsin L pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:23645]	-1.73566	0.000015
ENSG00000236417	CTSLP1	cathepsin L pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:2539]	-1.73566	0.000015
ENSG00000095627	TDRD1	tudor domain containing 1 [Source:HGNC Symbol;Acc:HGNC:11712]	2.85556	0.000016
ENSG00000213424	KRT222	keratin 222 [Source:HGNC Symbol;Acc:HGNC:28695]	1.4224	0.000052
ENSG00000237872	POU5F1P4	POU class 5 homeobox 1 pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:33310]	-1.52156	0.000102
ENSG00000128335	APOL2	apolipoprotein L2 [Source:HGNC Symbol;Acc:HGNC:619]	1.16331	0.000122
ENSG00000274818			-1.48522	0.000129
ENSG00000279364			-1.43321	0.000184
ENSG00000240694	PNMA2	paraneoplastic Ma antigen 2 [Source:HGNC Symbol;Acc:HGNC:9159]	1.12678	0.000207
ENSG00000111328	CDK2AP1	cyclin dependent kinase 2 associated protein 1 [Source:HGNC Symbol;Acc:HGNC:14002]	1.62401	0.000357
ENSG00000235423			1.62401	0.000357
ENSG00000105996	HOXA2	homeobox A2 [Source:HGNC Symbol;Acc:HGNC:5103]	1.17594	0.000363
ENSG00000169118	CSNK1G1	casein kinase 1 gamma 1 [Source:HGNC Symbol;Acc:HGNC:2454]	-1.07108	0.000367
ENSG00000178852	EFCAB13	EF-hand calcium binding domain 13 [Source:HGNC Symbol;Acc:HGNC:26864]	-1.06684	0.000396
ENSG00000183091	NEB	nebulin [Source:HGNC Symbol;Acc:HGNC:7720]	-1.124	0.000396
ENSG00000278373	G6PC2	glucose-6-phosphatase catalytic subunit 2 [Source:HGNC Symbol;Acc:HGNC:28906]	-1.21616	0.000417
ENSG00000225614	ZNF469	zinc finger protein 469 [Source:HGNC Symbol;Acc:HGNC:23216]	-1.39136	0.000428
ENSG00000228517	CT47A7	cancer/testis antigen family 47, member A7 [Source:HGNC Symbol;Acc:HGNC:33288]	-1.20097	0.000431
ENSG00000197893	NRAP	nebulin related anchoring protein [Source:HGNC Symbol;Acc:HGNC:7988]	1.07561	0.000445
ENSG00000270392	PFN1P2	profilin 1 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:24298]	1.32941	0.000468
ENSG00000236562			1.93558	0.00053
ENSG00000232559			-1.31108	0.00058
ENSG00000269936			-3.46921	0.000587
ENSG00000158352	SHROOM4	shroom family member 4 [Source:HGNC Symbol;Acc:HGNC:29215]	1.03569	0.000589
ENSG00000196116	TDRD7	tudor domain containing 7 [Source:HGNC Symbol;Acc:HGNC:30831]	1.03901	0.000592
ENSG00000265790	RNASEH1P1	ribonuclease H1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:10049]	-1.29018	0.000592
ENSG00000108830	RND2	Rho family GTPase 2 [Source:HGNC Symbol;Acc:HGNC:18315]	-1.03785	0.000618
ENSG00000278494	RRN3	RRN3 homolog, RNA polymerase I transcription factor [Source:HGNC Symbol;Acc:HGNC:30346]	-1.16786	0.000633
ENSG00000176945	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637
ENSG00000274843	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637
ENSG00000275501	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637

ENSG00000276583	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637
ENSG00000277753	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637
ENSG00000278114	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637
ENSG00000281630	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.41334	0.000637
ENSG00000198169	ZNF251	zinc finger protein 251 [Source:HGNC Symbol;Acc:HGNC:13045]	-1.21859	0.000655
ENSG00000275430	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]	-1.30138	0.000656
ENSG00000268942	CKS1BP3	CDC28 protein kinase regulatory subunit 1B pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:24233]	-1.28138	0.000672
ENSG00000156504	FAM122B	family with sequence similarity 122B [Source:HGNC Symbol;Acc:HGNC:30490]	-1.01605	0.000706
ENSG00000137509	PRCP	prolylcarboxypeptidase [Source:HGNC Symbol;Acc:HGNC:9344]	1.08231	0.000715
ENSG00000272841			-1.29057	0.000719
ENSG00000133169	BEX1	brain expressed X-linked 1 [Source:HGNC Symbol;Acc:HGNC:1036]	-1.02748	0.000723
ENSG00000243828			1.38453	0.000753
ENSG00000111834	RSPH4A	radial spoke head 4 homolog A [Source:HGNC Symbol;Acc:HGNC:21558]	-1.00468	0.00083
ENSG00000262202			-1.25291	0.000832
ENSG00000221968	FADS3	fatty acid desaturase 3 [Source:HGNC Symbol;Acc:HGNC:3576]	-1.00994	0.000836
ENSG00000155090	KLF10	Kruppel like factor 10 [Source:HGNC Symbol;Acc:HGNC:11810]	1.94444	0.000874
ENSG00000133624	ZNF767P	zinc finger family member 767, pseudogene [Source:HGNC Symbol;Acc:HGNC:21884]	-1.31353	0.000906
ENSG00000163686	ABHD6	abhydrolase domain containing 6 [Source:HGNC Symbol;Acc:HGNC:21398]	-0.98849	0.000909
ENSG00000181101	SDCCAG3P2	serologically defined colon cancer antigen 3 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:39164]	1.23517	0.000938
ENSG00000103249	CLCN7	chloride voltage-gated channel 7 [Source:HGNC Symbol;Acc:HGNC:2025]	-0.98984	0.000976
ENSG00000006607	FARP2	FERM, ARH/RhoGEF and pleckstrin domain protein 2 [Source:HGNC Symbol;Acc:HGNC:16460]	-1.01743	0.00105
ENSG00000004848	ARX	aristless related homeobox [Source:HGNC Symbol;Acc:HGNC:18060]	-1.03587	0.001104
ENSG00000196268	ZNF493	zinc finger protein 493 [Source:HGNC Symbol;Acc:HGNC:23708]	-1.0958	0.001107
ENSG00000239665			-1.19454	0.00123
ENSG00000135423	GLS2	glutaminase 2 [Source:HGNC Symbol;Acc:HGNC:29570]	-1.10166	0.001241
ENSG00000165512	ZNF22	zinc finger protein 22 [Source:HGNC Symbol;Acc:HGNC:13012]	0.964531	0.001269
ENSG00000215096	IFITM8P	interferon induced transmembrane protein 8 pseudogene [Source:HGNC Symbol;Acc:HGNC:32202]	1.21692	0.001284
ENSG00000179104	TMTC2	transmembrane and tetratricopeptide repeat containing 2 [Source:HGNC Symbol;Acc:HGNC:25440]	0.966941	0.001311
ENSG00000157020	SEC13	SEC13 homolog, nuclear pore and COPII coat complex component [Source:HGNC Symbol;Acc:HGNC:10697]	1.05032	0.001344
ENSG00000135222	CSN2	casein beta [Source:HGNC Symbol;Acc:HGNC:2447]	0.953723	0.001366
ENSG00000279738			-3.26814	0.001389
ENSG00000172890	NADSYN1	NAD synthetase 1 [Source:HGNC Symbol;Acc:HGNC:29832]	-1.20149	0.001419
ENSG00000095637	SORBS1	sorbin and SH3 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:14565]	0.949884	0.001434
ENSG00000121764	HCRTR1	hypocretin receptor 1 [Source:HGNC Symbol;Acc:HGNC:4848]	0.949082	0.001502

ENSG00000152078	TMEM56	transmembrane protein 56 [Source:HGNC Symbol;Acc:HGNC:26477]	1.06161	0.001514
ENSG00000181804	SLC9A9	solute carrier family 9 member A9 [Source:HGNC Symbol;Acc:HGNC:20653]	-0.94273	0.001594
ENSG00000204604	ZNF468	zinc finger protein 468 [Source:HGNC Symbol;Acc:HGNC:33105]	-1.0549	0.001601
ENSG00000267149			-1.17186	0.00161
ENSG00000274413			-1.17186	0.00161
ENSG00000203859	HSD3B2	hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid delta-isomerase 2 [Source:HGNC Symbol;Acc:HGNC:5218]	-0.93564	0.001644
ENSG00000231305			2.84756	0.001686
ENSG00000130684	ZNF337	zinc finger protein 337 [Source:HGNC Symbol;Acc:HGNC:15809]	-1.24861	0.001691
ENSG00000175147	TMEM51-AS1	TMEM51 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:26301]	-1.01445	0.001701
ENSG00000125869	LAMP5	lysosomal associated membrane protein family member 5 [Source:HGNC Symbol;Acc:HGNC:16097]	0.942166	0.001739
ENSG00000280318			-2.81294	0.001802
ENSG00000235574			-1.16094	0.00186
ENSG00000075975	MKRN2	makorin ring finger protein 2 [Source:HGNC Symbol;Acc:HGNC:7113]	1.78123	0.001881
ENSG00000103653	CSK	c-src tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:2444]	0.919041	0.001901
ENSG00000283030	CSN2	casein beta [Source:HGNC Symbol;Acc:HGNC:2447]	1.03618	0.001903
ENSG00000038274	MAT2B	methionine adenosyltransferase 2B [Source:HGNC Symbol;Acc:HGNC:6905]	1.1597	0.00192
ENSG00000198601	OR2M2	olfactory receptor family 2 subfamily M member 2 [Source:HGNC Symbol;Acc:HGNC:8268]	1.73823	0.00197
ENSG00000176700	SCAND2P	SCAN domain containing 2 pseudogene [Source:HGNC Symbol;Acc:HGNC:10567]	-1.07796	0.002004
ENSG00000255413			3.08198	0.002012
ENSG00000163364	LINC01116	long intergenic non-protein coding RNA 1116 [Source:HGNC Symbol;Acc:HGNC:49259]	1.1481	0.002065
ENSG00000213160	KLHL23	kelch like family member 23 [Source:HGNC Symbol;Acc:HGNC:27506]	0.921493	0.002096
ENSG00000237514	PTP4A1P7	protein tyrosine phosphatase type IVA, member 1 pseudogene 7 [Source:HGNC Symbol;Acc:HGNC:41934]	-1.15032	0.002134
ENSG00000175866	BAIAP2	BAI1 associated protein 2 [Source:HGNC Symbol;Acc:HGNC:947]	-1.29139	0.002233
ENSG00000237499			-1.13095	0.002234
ENSG00000142675	CNKSR1	connector enhancer of kinase suppressor of Ras 1 [Source:HGNC Symbol;Acc:HGNC:19700]	0.920822	0.002238
ENSG00000198155	ZNF876P	zinc finger protein 876, pseudogene [Source:HGNC Symbol;Acc:HGNC:32472]	-1.14257	0.002282
ENSG00000170298	LGALS9B	galectin 9B [Source:HGNC Symbol;Acc:HGNC:24842]	1.01997	0.002364
ENSG00000124209	RAB22A	RAB22A, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:9764]	-0.90329	0.002372
ENSG00000155256	ZFYVE27	zinc finger FYVE-type containing 27 [Source:HGNC Symbol;Acc:HGNC:26559]	-0.90882	0.002429
ENSG00000151445	VIPAS39	VPS33B interacting protein, apical-basolateral polarity regulator, spe-39 homolog [Source:HGNC Symbol;Acc:HGNC:20347]	1.22812	0.002539
ENSG00000279119			2.6364	0.002563
ENSG00000241322	CDRT1	CMT1A duplicated region transcript 1 [Source:HGNC Symbol;Acc:HGNC:14379]	-1.00685	0.002575
ENSG00000277564	RBFOX2	RNA binding protein, fox-1 homolog 2 [Source:HGNC Symbol;Acc:HGNC:9906]	1.00178	0.002647
ENSG00000152518	ZFP36L2	ZFP36 ring finger protein like 2 [Source:HGNC Symbol;Acc:HGNC:1108]	0.880635	0.002779
ENSG00000070019	GUCY2C	guanylate cyclase 2C [Source:HGNC Symbol;Acc:HGNC:4688]	0.880172	0.002795

ENSG00000107164	FUBP3	far upstream element binding protein 3 [Source:HGNC Symbol;Acc:HGNC:4005]	-1.02005	0.002802
ENSG00000104375	STK3	serine/threonine kinase 3 [Source:HGNC Symbol;Acc:HGNC:11406]	0.890932	0.002829
ENSG00000132780	NASP	nuclear autoantigenic sperm protein [Source:HGNC Symbol;Acc:HGNC:7644]	-0.88396	0.002858
ENSG00000206535	LNP1	leukemia NUP98 fusion partner 1 [Source:HGNC Symbol;Acc:HGNC:28014]	2.57907	0.002887
ENSG0000019549	SNAI2	snail family transcriptional repressor 2 [Source:HGNC Symbol;Acc:HGNC:11094]	0.925551	0.002904
ENSG00000184709	LRRC26	leucine rich repeat containing 26 [Source:HGNC Symbol;Acc:HGNC:31409]	-2.87893	0.003067
ENSG00000149300	C11orf52	chromosome 11 open reading frame 52 [Source:HGNC Symbol;Acc:HGNC:30531]	0.984652	0.003077
ENSG00000196152	ZNF79	zinc finger protein 79 [Source:HGNC Symbol;Acc:HGNC:13153]	-0.94805	0.003095
ENSG00000105376	ICAM5	intercellular adhesion molecule 5 [Source:HGNC Symbol;Acc:HGNC:5348]	-0.88807	0.00314
ENSG00000075151	EIF4G3	eukaryotic translation initiation factor 4 gamma 3 [Source:HGNC Symbol;Acc:HGNC:3298]	-0.90789	0.003174
ENSG00000277292	PIP4K2B	phosphatidylinositol-5-phosphate 4-kinase type 2 beta [Source:HGNC Symbol;Acc:HGNC:8998]	0.983993	0.003279
ENSG00000185885	IFITM1	interferon induced transmembrane protein 1 [Source:HGNC Symbol;Acc:HGNC:5412]	1.26129	0.003306
ENSG00000225914			2.82594	0.003433
ENSG00000137573	SULF1	sulfatase 1 [Source:HGNC Symbol;Acc:HGNC:20391]	0.870293	0.003482
ENSG00000100320	RBFOX2	RNA binding protein, fox-1 homolog 2 [Source:HGNC Symbol;Acc:HGNC:9906]	0.865211	0.003529
ENSG00000008405	CRY1	cryptochrome circadian clock 1 [Source:HGNC Symbol;Acc:HGNC:2384]	-0.86282	0.003579
ENSG00000143653	SCCPDH	saccharopine dehydrogenase (putative) [Source:HGNC Symbol;Acc:HGNC:24275]	-0.8591	0.003774
ENSG00000130147	SH3BP4	SH3 domain binding protein 4 [Source:HGNC Symbol;Acc:HGNC:10826]	0.858072	0.003802
ENSG00000263293		thyroid cancer-associated transcript 158 [Source:EntrezGene;Acc:102724508]	-1.1466	0.003851
ENSG00000072110	ACTN1	actinin alpha 1 [Source:HGNC Symbol;Acc:HGNC:163]	0.857799	0.003972
ENSG00000163106	HPGDS	hematopoietic prostaglandin D synthase [Source:HGNC Symbol;Acc:HGNC:17890]	-0.86042	0.004043
ENSG00000104808	DHDH	dihydrodiol dehydrogenase [Source:HGNC Symbol;Acc:HGNC:17887]	-0.92633	0.004061
ENSG00000131914	LIN28A	lin-28 homolog A [Source:HGNC Symbol;Acc:HGNC:15986]	0.846846	0.004094
ENSG00000138944	KIAA1644	KIAA1644 [Source:HGNC Symbol;Acc:HGNC:29335]	-0.92115	0.004094
ENSG00000171916	LGALS9C	galectin 9C [Source:HGNC Symbol;Acc:HGNC:33874]	0.960653	0.00411
ENSG00000163898	LIPH	lipase H [Source:HGNC Symbol;Acc:HGNC:18483]	-0.84099	0.004164
ENSG00000106615	RHEB	Ras homolog enriched in brain [Source:HGNC Symbol;Acc:HGNC:10011]	-0.83585	0.004238
ENSG00000214518	KRTAP2-2	keratin associated protein 2-2 [Source:HGNC Symbol;Acc:HGNC:18905]	-2.39611	0.004286
ENSG00000274887	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000274953	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000275330	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000275577	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000275962	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000276066	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000276159	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000276363	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351

ENSG00000277607	VSTM1	V-set and transmembrane domain containing 1 [Source:HGNC Symbol;Acc:HGNC:29455]	-2.7164	0.004351
ENSG00000228801			2.38253	0.004418
ENSG00000254445	HSPB2-C11orf52	HSPB2-C11orf52 readthrough (NMD candidate) [Source:HGNC Symbol;Acc:HGNC:41996]	0.94271	0.004423
ENSG00000169249	ZRSR2	zinc finger CCCH-type, RNA binding motif and serine/arginine rich 2 [Source:HGNC Symbol;Acc:HGNC:23019]	-0.8342	0.004455
ENSG00000187997	C17orf99	chromosome 17 open reading frame 99 [Source:HGNC Symbol;Acc:HGNC:34490]	-0.91418	0.00449
ENSG00000112299	VNN1	vanin 1 [Source:HGNC Symbol;Acc:HGNC:12705]	-0.84576	0.004531
ENSG00000165115	KIF27	kinesin family member 27 [Source:HGNC Symbol;Acc:HGNC:18632]	-0.93376	0.004628
ENSG00000231643			2.35138	0.004738
ENSG0000028310	BRD9	bromodomain containing 9 [Source:HGNC Symbol;Acc:HGNC:25818]	-0.83171	0.004745
ENSG00000110169	HPX	hemopexin [Source:HGNC Symbol;Acc:HGNC:5171]	-0.83931	0.004843
ENSG00000214954	LRRC69	leucine rich repeat containing 69 [Source:HGNC Symbol;Acc:HGNC:34303]	2.34073	0.004853
ENSG00000167916	KRT24	keratin 24 [Source:HGNC Symbol;Acc:HGNC:18527]	-1.21158	0.00487
ENSG00000124762	CDKN1A	cyclin dependent kinase inhibitor 1A [Source:HGNC Symbol;Acc:HGNC:1784]	-1.08511	0.004891
ENSG00000185619	PCGF3	polycomb group ring finger 3 [Source:HGNC Symbol;Acc:HGNC:10066]	-0.94984	0.005023
ENSG00000069424	KCNAB2	potassium voltage-gated channel subfamily A regulatory beta subunit 2 [Source:HGNC Symbol;Acc:HGNC:6229]	-0.87827	0.005121
ENSG00000204572	KRTAP5-10	keratin associated protein 5-10 [Source:HGNC Symbol;Acc:HGNC:23605]	-1.75974	0.005148
ENSG00000102387	TAF7L	TATA-box binding protein associated factor 7 like [Source:HGNC Symbol;Acc:HGNC:11548]	0.830012	0.00515
ENSG00000164111	ANXA5	annexin A5 [Source:HGNC Symbol;Acc:HGNC:543]	0.834164	0.005151
ENSG00000172362	OR5B12	olfactory receptor family 5 subfamily B member 12 [Source:HGNC Symbol;Acc:HGNC:15432]	1.53176	0.005202
ENSG00000255145	STX17-AS1	STX17 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:51174]	-1.02262	0.005224
ENSG00000170381	SEMA3E	semaphorin 3E [Source:HGNC Symbol;Acc:HGNC:10727]	0.829688	0.00523
ENSG00000233419			1.64714	0.005272
ENSG00000198960	ARMCX6	armadillo repeat containing, X-linked 6 [Source:HGNC Symbol;Acc:HGNC:26094]	0.930398	0.005281
ENSG00000184863	RBM33	RNA binding motif protein 33 [Source:HGNC Symbol;Acc:HGNC:27223]	-0.83098	0.005309
ENSG00000182952	HMGNA4	high mobility group nucleosomal binding domain 4 [Source:HGNC Symbol;Acc:HGNC:4989]	0.885762	0.005318
ENSG00000221819	GAS8-AS1	GAS8 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:1197]	0.822458	0.005325
ENSG00000165879	FRAT1	FRAT1, WNT signaling pathway regulator [Source:HGNC Symbol;Acc:HGNC:3944]	-1.52791	0.005327
ENSG00000166578	IQCD	IQ motif containing D [Source:HGNC Symbol;Acc:HGNC:25168]	-1.51111	0.005394
ENSG00000105664	COMP	cartilage oligomeric matrix protein [Source:HGNC Symbol;Acc:HGNC:2227]	-0.82405	0.005404
ENSG00000197617	VN1R5	vomer nasal 1 receptor 5 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:19870]	1.08025	0.005502
ENSG00000170236	USP50	ubiquitin specific peptidase 50 [Source:HGNC Symbol;Acc:HGNC:20079]	1.5155	0.005573
ENSG00000079841	RIMS1	regulating synaptic membrane exocytosis 1 [Source:HGNC Symbol;Acc:HGNC:17282]	0.830559	0.005628
ENSG00000132429	POPDC3	popeye domain containing 3 [Source:HGNC Symbol;Acc:HGNC:17649]	-1.13636	0.005672
ENSG00000189023	MAGEB16	MAGE family member B16 [Source:HGNC Symbol;Acc:HGNC:21188]	-2.27242	0.005672
ENSG00000188011	RTP5	receptor transporter protein 5 (putative) [Source:HGNC Symbol;Acc:HGNC:26585]	-0.81263	0.005687
ENSG00000254369	HOXA-AS3	HOXA cluster antisense RNA 3 [Source:HGNC Symbol;Acc:HGNC:43748]	1.01054	0.005715

ENSG00000198429	ZNF69	zinc finger protein 69 [Source:HGNC Symbol;Acc:HGNC:13138]	1.5094	0.005772
ENSG00000011052	NME1-NME2	NME1-NME2 readthrough [Source:HGNC Symbol;Acc:HGNC:33531]	1.31296	0.005797
ENSG00000169327	OR5AU1	olfactory receptor family 5 subfamily AU member 1 [Source:HGNC Symbol;Acc:HGNC:15362]	-1.52044	0.005935
ENSG00000111911	HINT3	histidine triad nucleotide binding protein 3 [Source:HGNC Symbol;Acc:HGNC:18468]	0.808889	0.005956
ENSG00000245694	CRNDE	colorectal neoplasia differentially expressed [Source:HGNC Symbol;Acc:HGNC:37078]	0.885372	0.005986
ENSG00000275867	KANSL1	KAT8 regulatory NSL complex subunit 1 [Source:HGNC Symbol;Acc:HGNC:24565]	0.905385	0.006052
ENSG00000278458	KANSL1	KAT8 regulatory NSL complex subunit 1 [Source:HGNC Symbol;Acc:HGNC:24565]	0.905385	0.006052
ENSG00000258297			-1.00064	0.006054
ENSG00000036530	CYP46A1	cytochrome P450 family 46 subfamily A member 1 [Source:HGNC Symbol;Acc:HGNC:2641]	0.811195	0.006084
ENSG00000132631	SCP2D1	SCP2 sterol binding domain containing 1 [Source:HGNC Symbol;Acc:HGNC:16211]	0.812277	0.00609
ENSG00000065802	ASB1	ankyrin repeat and SOCS box containing 1 [Source:HGNC Symbol;Acc:HGNC:16011]	-0.81521	0.006282
ENSG00000226400			2.2271	0.0063
ENSG00000180332	KCTD4	potassium channel tetramerization domain containing 4 [Source:HGNC Symbol;Acc:HGNC:23227]	0.807457	0.00633
ENSG00000262628	OR1D5	olfactory receptor family 1 subfamily D member 5 [Source:HGNC Symbol;Acc:HGNC:8186]	-0.87283	0.006341
ENSG00000152254	G6PC2	glucose-6-phosphatase catalytic subunit 2 [Source:HGNC Symbol;Acc:HGNC:28906]	-1.02018	0.00638
ENSG00000186185	KIF18B	kinesin family member 18B [Source:HGNC Symbol;Acc:HGNC:27102]	1.00156	0.00643
ENSG00000276293	PIP4K2B	phosphatidylinositol-5-phosphate 4-kinase type 2 beta [Source:HGNC Symbol;Acc:HGNC:8998]	0.809258	0.00647
ENSG00000203993	ARRDC1-AS1	ARRDC1 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:23395]	-1.07151	0.006505
ENSG00000172548	NIPAL4	NIPA like domain containing 4 [Source:HGNC Symbol;Acc:HGNC:28018]	-2.21043	0.006551
ENSG00000186432	KPNA4	karyopherin subunit alpha 4 [Source:HGNC Symbol;Acc:HGNC:6397]	-0.82673	0.006555
ENSG00000163684	RPP14	ribonuclease P/MRP subunit p14 [Source:HGNC Symbol;Acc:HGNC:30327]	-0.87319	0.006633
ENSG00000161249	DMKN	dermokine [Source:HGNC Symbol;Acc:HGNC:25063]	-1.09234	0.00669
ENSG00000279759			0.992683	0.006695
ENSG00000197343	ZNF655	zinc finger protein 655 [Source:HGNC Symbol;Acc:HGNC:30899]	-0.88952	0.00683
ENSG00000025772	TOMM34	translocase of outer mitochondrial membrane 34 [Source:HGNC Symbol;Acc:HGNC:15746]	-0.86865	0.006837
ENSG00000215102	TERF1P4	telomeric repeat binding factor 1 pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:38500]	2.50556	0.006948
ENSG00000160917	CPSF4	cleavage and polyadenylation specific factor 4 [Source:HGNC Symbol;Acc:HGNC:2327]	-1.10079	0.006949
ENSG00000083307	GRHL2	grainyhead like transcription factor 2 [Source:HGNC Symbol;Acc:HGNC:2799]	0.864783	0.006966
ENSG00000109919	MTCH2	mitochondrial carrier 2 [Source:HGNC Symbol;Acc:HGNC:17587]	-1.4947	0.007166
ENSG00000091947	TMEM101	transmembrane protein 101 [Source:HGNC Symbol;Acc:HGNC:28653]	0.79231	0.007174
ENSG00000278062	ZNF251	zinc finger protein 251 [Source:HGNC Symbol;Acc:HGNC:13045]	-0.98122	0.007209
ENSG00000130164	LDLR	low density lipoprotein receptor [Source:HGNC Symbol;Acc:HGNC:6547]	0.784163	0.007303
ENSG00000267350			-2.16427	0.007304
ENSG00000235978			-0.97675	0.007331
ENSG00000170293	CMTM8	CKLF like MARVEL transmembrane domain containing 8 [Source:HGNC Symbol;Acc:HGNC:19179]	0.792901	0.007371

ENSG00000144909	OSBPL11	oxysterol binding protein like 11 [Source:HGNC Symbol;Acc:HGNC:16397]	0.792443	0.007393
ENSG00000093167	LRRFIP2	LRR binding FLII interacting protein 2 [Source:HGNC Symbol;Acc:HGNC:6703]	-0.78635	0.007397
ENSG00000108091	CCDC6	coiled-coil domain containing 6 [Source:HGNC Symbol;Acc:HGNC:18782]	-1.45322	0.0074
ENSG00000125484	GTF3C4	general transcription factor IIIC subunit 4 [Source:HGNC Symbol;Acc:HGNC:4667]	0.892655	0.007461
ENSG00000103550	KNOP1	lysine rich nucleolar protein 1 [Source:HGNC Symbol;Acc:HGNC:34404]	1.03209	0.00752
ENSG00000227940	LINC01696	long intergenic non-protein coding RNA 1696 [Source:HGNC Symbol;Acc:HGNC:52484]	-2.15107	0.007537
ENSG00000178950	GAK	cyclin G associated kinase [Source:HGNC Symbol;Acc:HGNC:4113]	-0.78168	0.00759
ENSG00000183840	GPR39	G protein-coupled receptor 39 [Source:HGNC Symbol;Acc:HGNC:4496]	0.795614	0.007652
ENSG00000252690	SCARNA15	small Cajal body-specific RNA 15 [Source:HGNC Symbol;Acc:HGNC:32572]	-1.30536	0.00768
ENSG00000204072	ARMCX7P	armadillo repeat containing, X-linked 7, pseudogene [Source:HGNC Symbol;Acc:HGNC:49917]	0.885434	0.007749
ENSG00000110076	NRXN2	neurexin 2 [Source:HGNC Symbol;Acc:HGNC:8009]	-0.78298	0.007753
ENSG00000259051	HNRNPUP1	heterogeneous nuclear ribonucleoprotein U pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:19931]	-0.97671	0.00789
ENSG00000277877			-0.97671	0.00789
ENSG00000090013	BLVRB	biliverdin reductase B [Source:HGNC Symbol;Acc:HGNC:1063]	-0.85564	0.007972
ENSG00000138459	SLC35A5	solute carrier family 35 member A5 [Source:HGNC Symbol;Acc:HGNC:20792]	0.786169	0.008142
ENSG00000070882	OSBPL3	oxysterol binding protein like 3 [Source:HGNC Symbol;Acc:HGNC:16370]	-1.46119	0.008229
ENSG00000180806	HOXC9	homeobox C9 [Source:HGNC Symbol;Acc:HGNC:5130]	0.861936	0.008277
ENSG00000186364	NUDT17	nudix hydrolase 17 [Source:HGNC Symbol;Acc:HGNC:26618]	1.44701	0.008385
ENSG00000256614	AK6P1	adenylate kinase 6 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:51678]	-0.96329	0.00852
ENSG00000260034	LCMT1-AS2	LCMT1 antisense RNA 2 [Source:HGNC Symbol;Acc:HGNC:51178]	-0.96453	0.008601
ENSG00000164920	OSR2	odd-skipped related transcription factor 2 [Source:HGNC Symbol;Acc:HGNC:15830]	0.781344	0.008607
ENSG00000110375	UPK2	uroplakin 2 [Source:HGNC Symbol;Acc:HGNC:12579]	0.802822	0.008611
ENSG00000106479	ZNF862	zinc finger protein 862 [Source:HGNC Symbol;Acc:HGNC:34519]	-0.82631	0.008624
ENSG00000121903	ZSCAN20	zinc finger and SCAN domain containing 20 [Source:HGNC Symbol;Acc:HGNC:13093]	-0.80983	0.008637
ENSG00000139910	NOVA1	NOVA alternative splicing regulator 1 [Source:HGNC Symbol;Acc:HGNC:7886]	0.877518	0.008671
ENSG00000108797	CNTNAP1	contactin associated protein 1 [Source:HGNC Symbol;Acc:HGNC:8011]	-1.45104	0.00871
ENSG00000161381	PLXDC1	plexin domain containing 1 [Source:HGNC Symbol;Acc:HGNC:20945]	-0.77596	0.008777
ENSG00000136932	TRMO	tRNA methyltransferase O [Source:HGNC Symbol;Acc:HGNC:30967]	-0.77108	0.008888
ENSG00000135480	KRT7	keratin 7 [Source:HGNC Symbol;Acc:HGNC:6445]	-0.76857	0.008931
ENSG00000213337	ANKRD39	ankyrin repeat domain 39 [Source:HGNC Symbol;Acc:HGNC:28640]	-0.97528	0.008941
ENSG00000130702	LAMA5	laminin subunit alpha 5 [Source:HGNC Symbol;Acc:HGNC:6485]	0.765189	0.008964
ENSG00000229989	MIR181A1HG	MIR181A1 host gene [Source:HGNC Symbol;Acc:HGNC:48659]	-2.07661	0.009013
ENSG00000187601	MAGEH1	MAGE family member H1 [Source:HGNC Symbol;Acc:HGNC:24092]	0.890517	0.009056
ENSG00000122547	EEPD1	endonuclease/exonuclease/phosphatase family domain containing 1 [Source:HGNC Symbol;Acc:HGNC:22223]	0.848997	0.009081
ENSG00000103647	CORO2B	coronin 2B [Source:HGNC Symbol;Acc:HGNC:2256]	-0.81859	0.009136

ENSG00000157502	MUM1L1	MUM1 like 1 [Source:HGNC Symbol;Acc:HGNC:26583]	-1.0577	0.00916
ENSG0000024048	UBR2	ubiquitin protein ligase E3 component n-recognin 2 [Source:HGNC Symbol;Acc:HGNC:21289]	0.873122	0.009169
ENSG00000183753	BPY2	basic charge, Y-linked, 2 [Source:HGNC Symbol;Acc:HGNC:13508]	0.855247	0.009189
ENSG00000183795	BPY2B	basic charge, Y-linked, 2B [Source:HGNC Symbol;Acc:HGNC:25449]	0.855247	0.009189
ENSG00000185894	BPY2C	basic charge, Y-linked, 2C [Source:HGNC Symbol;Acc:HGNC:18225]	0.855247	0.009189
ENSG00000169714	CNBP	CCHC-type zinc finger nucleic acid binding protein [Source:HGNC Symbol;Acc:HGNC:13164]	1.40533	0.009237
ENSG00000189410	SH2D5	SH2 domain containing 5 [Source:HGNC Symbol;Acc:HGNC:28819]	-0.82433	0.009242
ENSG00000247092	SNHG10	small nucleolar RNA host gene 10 [Source:HGNC Symbol;Acc:HGNC:27510]	-0.8544	0.00928
ENSG00000213204			-0.92343	0.009497
ENSG00000266970			-1.1573	0.009499
ENSG00000204946	ZNF783	zinc finger family member 783 [Source:HGNC Symbol;Acc:HGNC:27222]	-0.94489	0.009537
ENSG00000119689	DLST	dihydrolipoamide S-succinyltransferase [Source:HGNC Symbol;Acc:HGNC:2911]	1.00903	0.009632
ENSG00000171877	FRMD5	FERM domain containing 5 [Source:HGNC Symbol;Acc:HGNC:28214]	0.842093	0.009637
ENSG00000125571	IL37	interleukin 37 [Source:HGNC Symbol;Acc:HGNC:15563]	0.991014	0.009652
ENSG00000204689	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000206474	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000220550	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000224234	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000229412	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000230505	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000232397	OR10C1	olfactory receptor family 10 subfamily C member 1 (gene/pseudogene) [Source:HGNC Symbol;Acc:HGNC:8165]	-0.84473	0.009816
ENSG00000279941			-0.84473	0.009816

6.3 WebGestalt

Η λειτουργική ανάλυση εμπλουτισμού αποτελεί βασικό στάδιο για την ερμηνεία των συνόλων των γονιδίων που προέκυψαν από την ανάλυση αυτή. Για τον σκοπό αυτό θα χρησιμοποιηθεί το WEB-based Gene SeT AnaLysis Toolkit που αποτελεί ένα ολοκληρωμένο σύστημα εξόρυξης δεδομένων βασισμένο στον διαδικτυακό χώρο. Βασικές ενότητες αυτού αποτελούν η διαχείριση συνόλων γονιδίων, η ανάκτηση πληροφοριών, η ταξινόμηση, η οπτικοποίηση και τέλος οι διάφορες στατιστικές μέθοδοι (Wang et al., 2013; Zhang et al., 2005). Πριν επιδιώξουμε οποιαδήποτε ανάλυση σημειώθηκαν τα ENSG IDs (Yates et al., 2016) που αντιστοιχήθηκαν μοναδικά σε Entrez Gene IDs με τα οποία και λειτουργεί το σύστημα (Πίνακας 10). Ως γονίδια αναφοράς χρησιμοποιήθηκε το σύνολο των γονιδίων από κάθε ομάδα όπως εξήχθη από την μετα-ανάλυση, αγνοώντας τα γονίδια που εμφανίζονταν μόνο σε μία μελέτη. Στη συνέχεια δοκιμάστηκαν έξι διαφορετικές αναλύσεις από τις οποίες με διαφορετικούς συντελεστές προσαρμογής προέκυψαν τα αποτελέσματά μας (Πίνακας 11).

- **Gene Ontology (GO) Analysis** (Πίνακας 12, Πίνακας 18, Πίνακας 19, Πίνακας 20): μία από τις κύριες χρήσεις του GO (Ashburner et al., 2000; Gene Ontology, 2015) είναι να πραγματοποιήσει ανάλυση εμπλουτισμού σε σύνολα γονιδίων χρησιμοποιώντας αναφορές για τα σύνολα αυτά. Αυτή η υπηρεσία συνδέεται με το εργαλείο ανάλυσης από το σύστημα ταξινόμησης PANTHER (Mi et al., 2013).
- **Kyoto Encyclopedia of Genes and Genomes (KEGG) Analysis** (Πίνακας 13, Πίνακας 23): βασίζεται στην βάση δεδομένων KEGG, για την κατανόηση των λειτουργιών και της χρησιμότητας ενός βιολογικού συστήματος, όπως το κύτταρο, ο οργανισμός και το οικοσύστημα. Συμπεριλαμβάνει πληροφορίες μοριακού επιπέδου, μεγάλα σύνολα δεδομένων που παράγονται από την αλληλούχηση γονιδιώματος και άλλες πειραματικές τεχνολογίες (Kanehisa and Goto, 2000).
- **Pathway Commons (PC)** (Πίνακας 14, Πίνακας 22): βασίζεται σε μια συλλογή διαθέσιμων στο κοινό δεδομένων μονοπατιών από πολλαπλούς οργανισμούς. Τα μονοπάτια αυτά περιλαμβάνουν βιοχημικές αντιδράσεις, σύνθετα γεγονότα συναρμολόγησης, μεταφοράς και καταλύσεως και φυσικές αλληλεπιδράσεις που περιλαμβάνουν πρωτεΐνες, DNA, RNA, μικρά μόρια και σύμπλοκα (Cerami et al., 2011).

- **Wikipathways** (Πίνακας 15, Πίνακας 21): αποτελεί ανάλυση βασισμένη σε μια Wiki-based πηγή επεξεργασίας και επιμέλειας μονοπατιών, σε συνδυασμό με ένα ενσωματωμένο γραφικό εργαλείο επεξεργασίας βιολογικών μονοπατιών. Σκοπός της είναι η ανταπόκριση στην αυξανόμενη πρόκληση που παρουσιάζει η συνεχής εισροή βιολογικών δεδομένων και η παροχή ενός καινοτόμου παραδείγματος επιμέλειας περιεχομένου από τη βιολογική κοινότητα (Kelder et al., 2012; Pico et al., 2008).
- **Transcription Factor Target (TFT) Analysis** (Πίνακας 16, Πίνακας 24, Πίνακας 25): κατά την ανάλυση αυτή προκύπτουν μεταγραφικοί παράγοντες που αντιστοιχούν σε σύνολα γονιδίων. Η ανάλυση γίνεται ξεχωριστά για τα υπό και υπέρ-εκφρασμένα γονίδια. Η βάση δεδομένων για τους μεταγραφικούς παράγοντες και τα σύνολα γονιδίων είναι η MSigDB (Liberzon et al., 2015).
- **MicroRNA Analysis** (Πίνακας 17, Πίνακας 26, Πίνακας 27): κατά την ανάλυση αυτή προκύπτουν τα microRNAs που αντιστοιχούν σε γονίδια από αυτά που προέκυψαν από την ανάλυση αυτή. Γίνεται επίσης ξεχωριστά για τα υπό και υπέρ-εκφρασμένα γονίδια. Βάσεις δεδομένων για τα miR και τα γονίδια στόχους είναι οι MicroRNA Target όπως η mirDB (Wang, 2008).

Πίνακας 10: Αριθμός γονιδίων που αντιστοιχήθηκαν μοναδικά από εκείνα που εξήγαμε από την μετα-ανάλυση και χρησιμοποιήθηκαν από το σύστημα για τις αναλύσεις για κάθε σύνολο γονιδίων.

Group	Unique_Reference	Unique_Results
CH4_16	18087/26401	217/571
CL4_16	18342/26942	170/266

Πίνακας 11: Αναλύσεις που έδωσαν αποτελέσματα για κάθε σύνολο γονιδίων. Στις παρενθέσεις αναγράφεται το επίπεδο σημαντικότητας και ο αστερίσκος σηματοδοτεί την απουσία διόρθωσης πολλαπλών ερωτημάτων. Τέλος οι αναλύσεις TFT και microRNA έγιναν αφού διαχωρίσαμε τα up και down regulated γονίδια.

Group	GO	KEGG	Pathway Commons	Wiki Pathways	Transcription Factor Target	MicroRNA
CH4_16	(0.05)	(0.1)	(0.05)	(0.05*)	Down (0.05*)	Down (0.05*)
CL4_16	(0.1)	(0.05)	(0.05*)	(0.05*)	Up_Down(0.05*)	Up_Down(0.05*)

Όλες οι αναλύσεις βασίστηκαν στατιστικά στην υπεργεωμετρική κατανομή (Zhou, 1980). Η υπεργεωμετρική κατανομή αποτελεί μια διακριτή κατανομή πιθανότητας που περιγράφει την πιθανότητα k επιτυχιών σε n δοκιμές, χωρίς αντικατάσταση, από ένα πεπερασμένο πληθυσμό μεγέθους N που περιέχει ακριβώς K επιτυχίες, όπου κάθε δοκιμή θεωρείται είτε επιτυχία είτε αποτυχία. Στη στατιστική χρησιμοποιείται για τον υπολογισμό της στατιστικής σημασίας k επιτυχιών (από n συνολικές δοκιμές) από τον προαναφερθέντα πληθυσμό. Εφαρμόζεται συχνά για τον εντοπισμό των

υποπληθυσμών οι οποίοι υπερ- ή υπο-εκπροσωπούνται σε ένα δείγμα. Ως μέθοδος προσαρμογής πολλαπλών δοκιμών (*multiple test adjustment*) χρησιμοποιήθηκε η FDR (*false discovery rate*), ως μέθοδος ελέγχου για τις αναμενόμενες αναλογίες «θετικών» αποτελεσμάτων που στην πραγματικότητα είναι ψευδή (Benjamini and Hochberg, 1995; Zhang et al., 2005).

- **Αποτελέσματα σύγκρισης Control-High**

Πίνακας 12:Κυριότερα αποτελέσματα ανάλυσης GO για την σύγκριση control-high (Εικόνα 30).

Biological Process	#	Genes	GO ID	Statistics
DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	5	MDM2 CDKN1A UBC TRIAP1 PSMD12	0006977	rawP=0.0003 adjP=0.0161
signal transduction involved in mitotic cell cycle G1/S transition DNA damage checkpoint	5	MDM2 CDKN1A UBC TRIAP1 PSMD12	0072431	rawP=0.0003 adjP=0.0161
L-serine catabolic process	2	SHMT1 SDS	0006565	rawP=0.0002 adjP=0.0161

Πίνακας 13: Αποτελέσματα ανάλυσης KEGG για την σύγκριση control-high (Εικόνα 31, Εικόνα 32).

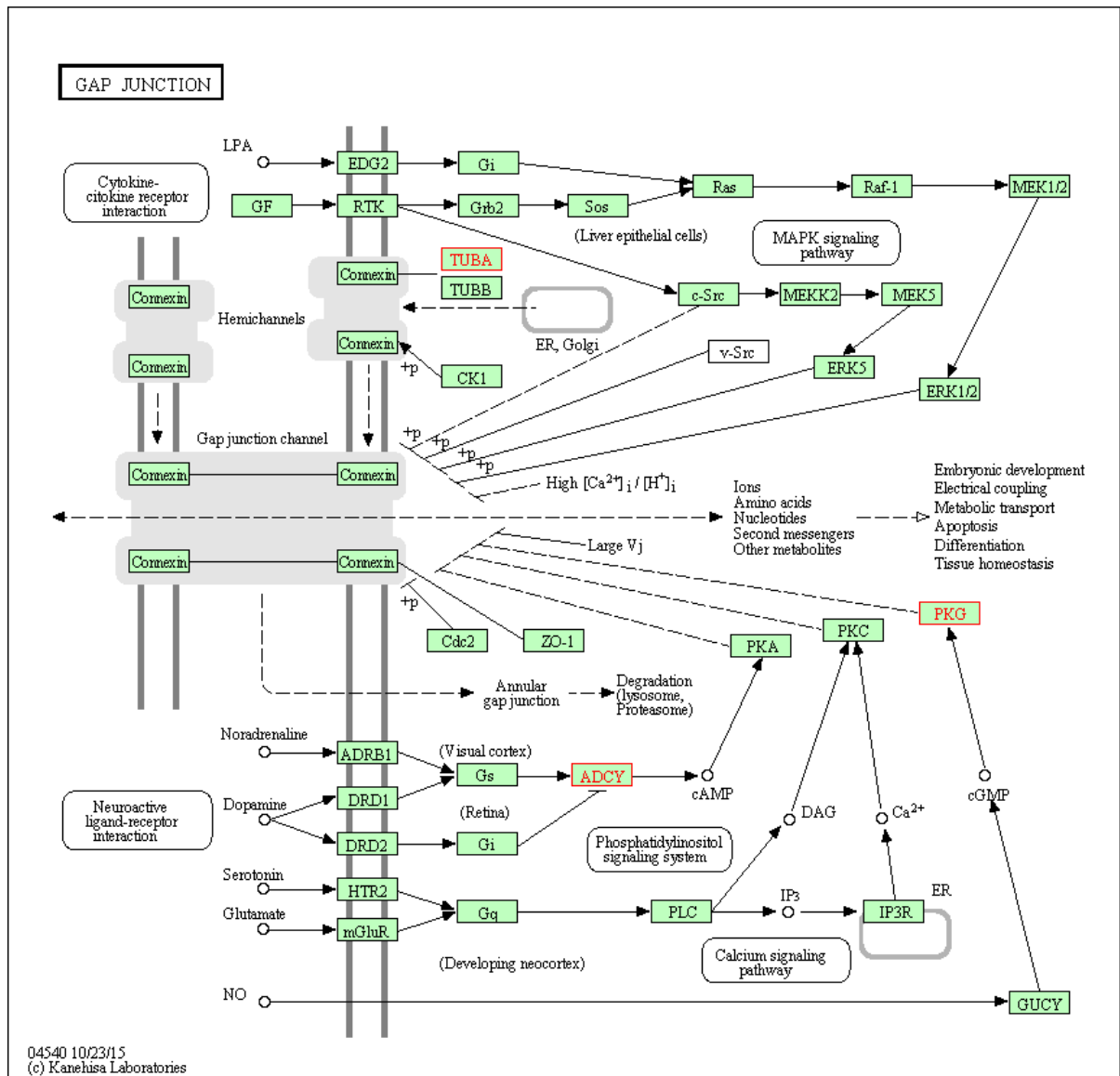
Pathway	#	Genes	ID	Statistics
Gap junction	5	PRKX TUBA1C ADCY4 ADCY9 PRKG1	04540	rawP=0.0016 adjP=0.0752
p53 signaling pathway	4	CCNG1 MDM2 CDKN1A SESN1	04115	rawP=0.0041 adjP=0.0964

Πίνακας 14: Αποτελέσματα ανάλυσης Pathway Commons για την σύγκριση control-high.

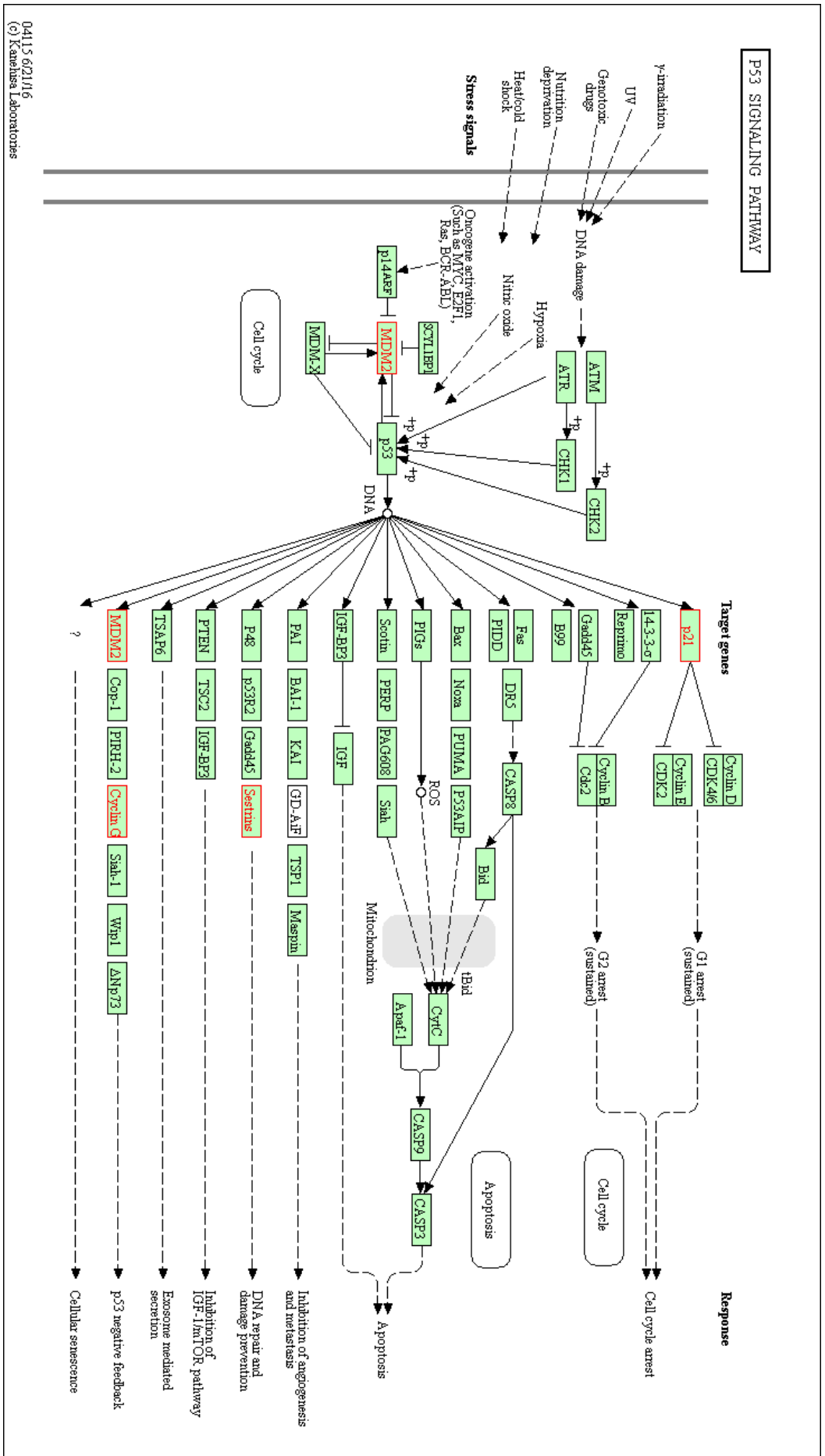
Pathway	#	Genes	ID	Statistics
p53 pathway	9	GDF15 CCNG1 MDM2 CDKN1A RPS27L C12orf5 SESN1 CSNK1G1 TRIAP1	1601	rawP=8.73e-05 adjP=0.0091
Direct p53 effectors	8	GDF15 CCNG1 MDM2 CDKN1A RPS27L C12orf5 SESN1 TRIAP1	1629	rawP=5.50e-05 adjP=0.0091
Signaling by EGFR	6	MDM2 CDKN1A ADCY4 ARHGEF7 UBC ADCY9	330	rawP=0.0004 adjP=0.0279
ATR signaling pathway	9	GDF15 CCNG1 MDM2 CDKN1A RPS27L C12orf5 SESN1 CSNK1G1 TRIAP1	1609	rawP=0.0007 adjP=0.0366

Πίνακας 15: Αποτελέσματα ανάλυσης Wiki Pathways για την σύγκριση control-high.

Pathway	#	Genes	ID	Statistics
DNA damage response (only ATM dependent)	3	CDKN1A FRAT1 LDLR	WP710	rawP=0.0488 adjP=0.0488



Εικόνα 31: Αναπαράσταση του Gap Junction pathway που προέκυψε από την ανάλυση KEGG για την σύγκριση control-high.



Εικόνα 32: Αναπαράσταση του p53 signaling pathway που προέκυψε από την ανάλυση KEGG για την σύγκριση control-high.

Πίνακας 16: Αποτελέσματα ανάλυσης Transcription Factor Target για την σύγκριση control-high για τα down-regulated γονίδια.

Target	#	Genes	ID	Statistics
hsa_CGTSACG_V\$PAX3_B	6	PCNP MRRF AFF4 CDKN1A EGR2 GLYR1	2444	rawP=0.0068 adjP=0.0068
hsa_V\$FREAC3_01	7	CDKN1A RUNX1T1 SLC35A2 ICAM5 SCRN3 EGR2 MAML3	2045	rawP=0.0249 adjP=0.0249

Πίνακας 17: Αποτελέσματα ανάλυσης microRNA Target για την σύγκριση control-high για τα down-regulated γονίδια

Target	#	Genes	ID	Statistics
hsa_ACACTAC MIR-142-3P	5	AFF1 PCGF3 PRLR EGR2 ADCY9	859	rawP=0.0174 adjP=0.0174
hsa_CCATCCA MIR-432	3	PCGF3 SLC7A8 PRKG1	783	rawP=0.0276 adjP=0.0276
hsa_TGTATGA MIR-485-3P	5	AFF4 PCGF3 GLS2 RUNX1T1 MAML3	735	rawP=0.0320 adjP=0.0320

- **Αποτελέσματα σύγκρισης Control-Low**

Πίνακας 18: Κυριότερα αποτελέσματα ανάλυσης GO για την σύγκριση control-low σχετικά με τα μέρη του κυττάρου (Εικόνα 33).

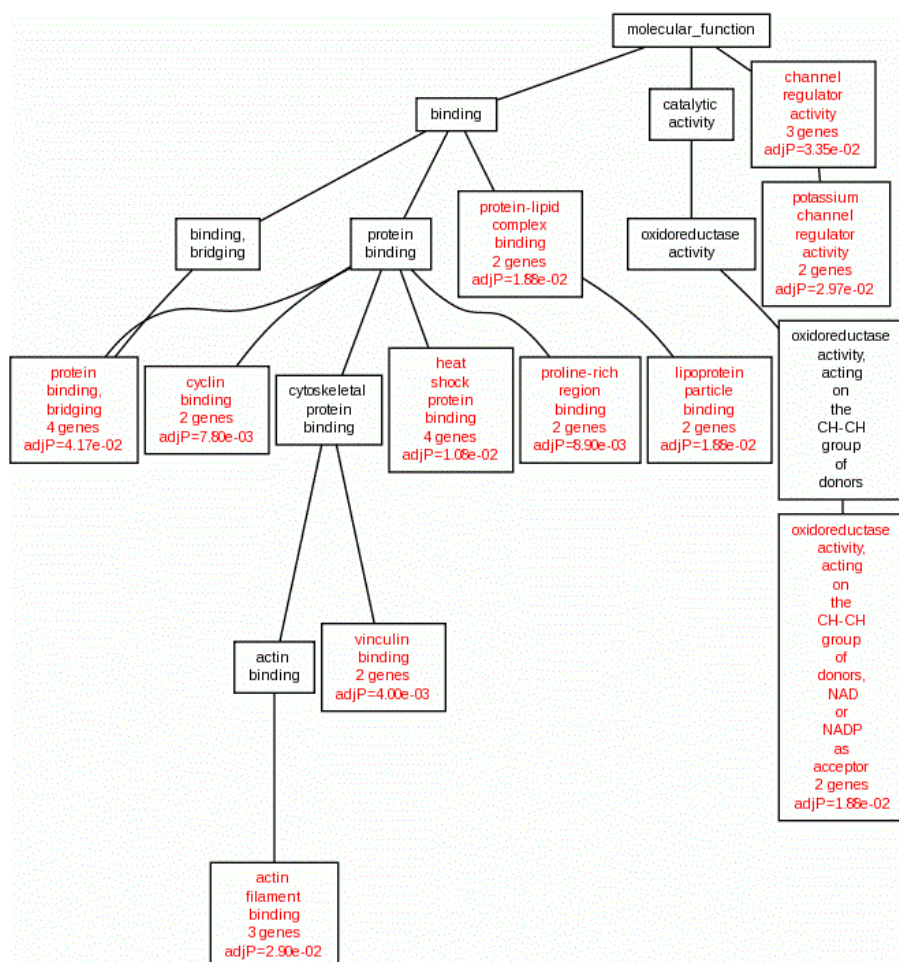
Cellular Component	#	Genes	GO ID	Statistics
Fascia Adherens	2	ACTN1 NRAP	0005916	rawP=0.0059 adjP=0.0915
Intercalated Disc	3	ANXA5 ACTN1 NRAP	0014704	rawP=0.0032 adjP=0.0915
Cell-Cell Contact Zone	3	ANXA5 ACTN1 NRAP	0044291	rawP=0.0038 adjP=0.0915
Recycling Endosome	5	SLC9A9 VIPAS39 ZFYVE27 LAMP5 LDLR	0055037	rawP=0.0004 adjP=0.0620
Growth Cone Membrane	2	ZFYVE27 LAMP5	0032584	rawP=0.0022 adjP=0.0915
Pole Plasm	2	TRD1 TRD7	0045495	rawP=0.0059 adjP=0.0915
Germ Plasm	2	TRD1 TRD7	0060293	rawP=0.0059 adjP=0.0915
P Granule	2	TRD1 TRD7	0043186	rawP=0.0059 adjP=0.0915
Chromatoid Body	2	TRD1 TRD7	0033391	rawP=0.0028 adjP=0.0915

Πίνακας 19: Κυριότερα αποτελέσματα ανάλυσης GO για την σύγκριση control-low χωρίς διόρθωση πολλαπλών ερωτημάτων στο κομμάτι των βιολογικών διεργασιών (Εικόνα 34).

Biological Process	#	Genes	GO ID	Statistics
focal adhesion assembly	3	ACTN1 SORBS1 LAMA5	0048041	rawP=0.0069 adjP=0.0069
protein localization to membrane	4	ZFYVE27 SORBS1 LAMA5 NRXN2	0072657	rawP=0.0109 adjP=0.0109
cell-substrate junction assembly	3	ACTN1 SORBS1 LAMA5	0007044	rawP=0.0159 adjP=0.0159
embryonic skeletal system development	4	SULF1 HOXC9 OSR2 HOXA2	0048706	rawP=0.0165 adjP=0.0165
middle ear morphogenesis	2	OSR2 HOXA2	0042474	rawP=0.0175 adjP=0.0175
cell junction assembly	5	CNTNAP1 ACTN1 SORBS1 LAMA5 SNAI2	0034329	rawP=0.0197 adjP=0.0197
NAD metabolic process	2	NADSYN1 DLST	0019674	rawP=0.0240 adjP=0.0240
cellular response to radiation	3	CDKN1A CRY1 SNAI2	0071478	rawP=0.0253 adjP=0.0253
cellular response to ionizing radiation	2	CDKN1A SNAI2	0071479	rawP=0.0374 adjP=0.0374
gene silencing	3	TDRD1 UBR2 LIN28A	0016458	rawP=0.0407 adjP=0.0407

Πίνακας 20: Κυριότερα αποτελέσματα ανάλυσης GO για την σύγκριση control-low χωρίς διόρθωση πολλαπλών ερωτημάτων στο κομμάτι των μοριακών λειτουργιών (Εικόνα 35).

Moleccular Function	#	Genes	GO ID	Statistics
vinculin binding	2	ACTN1 NRAP	0017166	rawP=0.0040 adjP=0.0040
cyclin binding	2	GAK CDKN1A	0030332	rawP=0.0078 adjP=0.0078
proline-rich region binding	2	CSK BAIAP2	0070064	rawP=0.0089 adjP=0.0089
heat shock protein binding	4	GAK NASP DLST TOMM34	0031072	rawP=0.0108 adjP=0.0108
oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	2	DHDH BLVRB	0016628	rawP=0.0188 adjP=0.0188
lipoprotein particle binding	2	APOL2 LDLR	0071813	rawP=0.0188 adjP=0.0188
protein-lipid complex binding	2	APOL2 LDLR	0071814	rawP=0.0188 adjP=0.0188



Εικόνα 35: Δίκτυο υπερεκπροσώπησης όρων γονιδιακής οντολογίας και ο συσχετισμός αυτών από την ανάλυση GO για την σύγκριση control-low χωρίς διόρθωση πολλαπλών ερωτημάτων. Τμήμα διαγράμματος για μοριακές λειτουργίες.

Πίνακας 21: Αποτελέσματα ανάλυσης Wiki Pathways για την σύγκριση control-low.

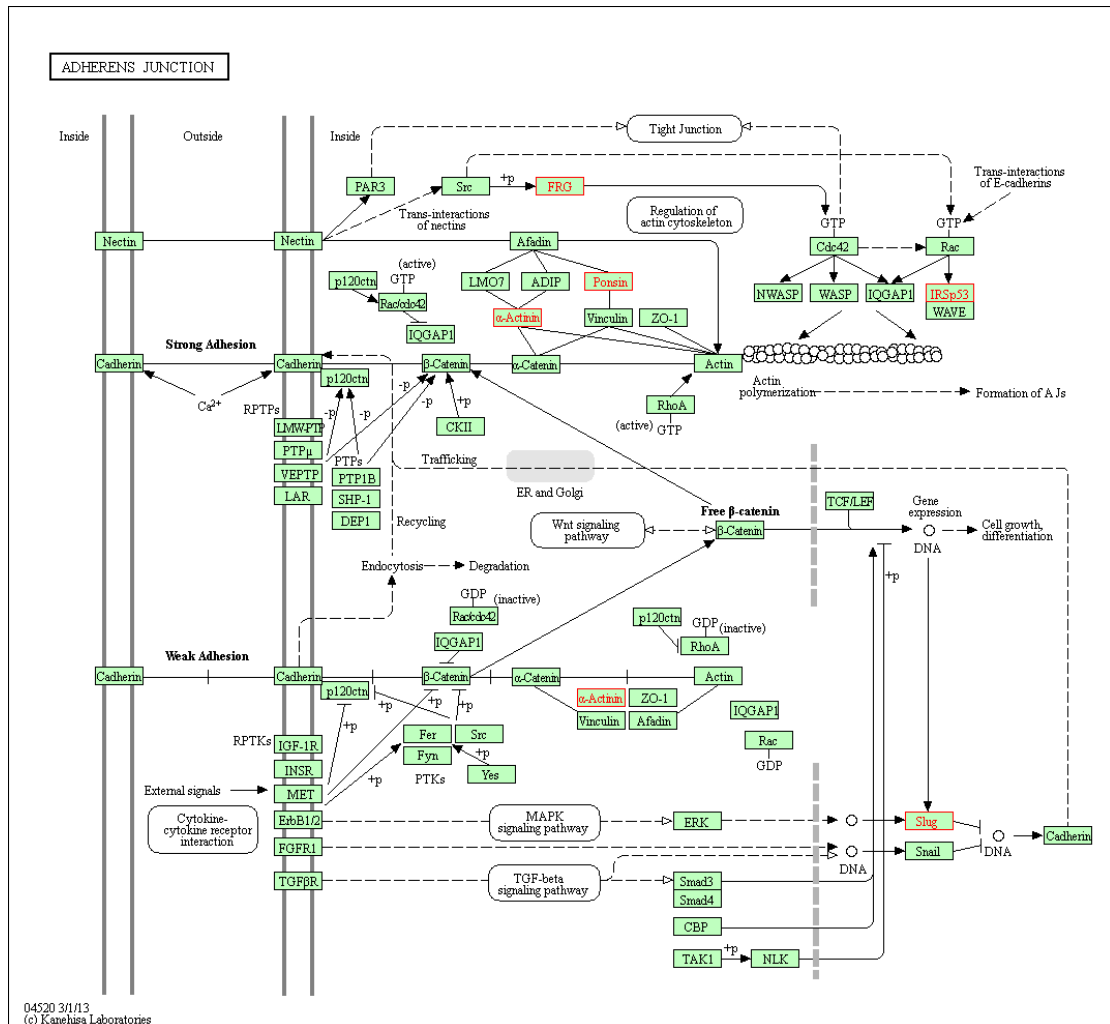
Pathway	#	Genes	ID	Statistics
DNA damage response (only ATM dependent)	3	CDKN1A FRAT1 LDLR	WP710	rawP=0.0498 adjP=0.0498

Πίνακας 22: Αποτελέσματα ανάλυσης Pathway Commons για την σύγκριση control-low.

Pathway	#	Genes	ID	Statistics
GAB1 signalosome	2	CDKN1A CSK	308	rawP=0.0353 adjP=0.0353

Πίνακας 23: Αποτελέσματα ανάλυσης KEGG για την σύγκριση control-low (Εικόνα 36).

Pathway	#	Genes	ID	Statistics
Adherens junction	5	ACTN1 SORBS1 BAIAP2 SNAI2 FARP2	04520	rawP=0.0005 adjP=0.0070



Εικόνα 36: Αναπαράσταση του Adherens Junction που προέκυψε από την ανάλυση KEGG για την σύγκριση control-low.

Πίνακας 24: Κυριότερα αποτελέσματα ανάλυσης Transcription Factor Target για την σύγκριση control-low για τα down-regulated γονίδια.

Target	#	Genes	ID	Statistics
hsa_V\$ROAZ_01	2	BEX1 CDKN1A	2121	rawP=0.0011 adjP=0.0011
hsa_V\$TCF11MAFG_01	4	POPDC3 LRRFIP2 NRXN2 BLVRB	2040	rawP=0.0181 adjP=0.0181

Πίνακας 25: Κυριότερα αποτελέσματα ανάλυσης Transcription Factor Target για την σύγκριση control-low για τα up-regulated γονίδια.

Target	#	Genes	ID	Statistics
hsa_V\$DR3_Q4	4	MAGEH1 ARMCX6 FRMD5 LIN28A	2330	rawP=0.0026 adjP=0.0026
hsa_V\$SOX5_01	5	ARMCX6 GRHL2 NOVA1 HOXA2 RBFOX2	1868	rawP=0.0047 adjP=0.0047
hsa_V\$MEF2_04	2	CYP46A1 NRAP	2009	rawP=0.0052 adjP=0.0052
hsa_CTGYNNCTYTAA _UNKNOWN	3	KRT222 OSR2 LIN28A	2159	rawP=0.0053 adjP=0.0053
hsa_TAATTA_V\$CHX10_01	9	KRT222 STK3 HOXA2 RBFOX2 LIN28A SULF1 OSR2 TMTC2 NOVA1	2408	rawP=0.0055 adjP=0.0055
hsa_V\$LXR_DR4_Q3	3	SULF1 FRMD5 LIN28A	2251	rawP=0.0063 adjP=0.0063

Πίνακας 26: Κυριότερα αποτελέσματα ανάλυσης microRNA Target για την σύγκριση control-low για τα down-regulated γονίδια.

Target	#	Genes	ID	Statistics
hsa_GTAGGCA MIR-189	2	KPNA4 RBM33	883	rawP=0.0085 adjP=0.0085
hsa_TTTGCAG MIR-518A-2	4	EIF4G3 KPNA4 RBM33 RAB22A	818	rawP=0.0184 adjP=0.0184
hsa_TGAATGT MIR-181A MIR-181B MIR-181C MIR-181D	6	FAM122B ZNF468 CSNK1G1 OSBPL3 ZNF655 BAIAP2	669	rawP=0.0294 adjP=0.0294

Πίνακας 27: Κυριότερα αποτελέσματα ανάλυσης microRNA Target για την σύγκριση control-low για τα up-regulated γονίδια.

Target	#	Genes	ID	Statistics
hsa_TTTGCAG MIR-518A-2	4	GUCY2C SORBS1 CDK2AP1 NOVA1	818	rawP=0.0115 adjP=0.0115
hsa_GCACCTT MIR-18A MIR-18B	3	GRHL2 LIN28A SH3BP4	708	rawP=0.0134 adjP=0.0134

6.4 Απεικόνιση Βιολογικών Δικτύων (STRING)

Οι πολλαπλές λειτουργικές αλληλεπιδράσεις και συνεργασίες που συμβαίνουν μεταξύ πρωτεϊνών αποτελούν τον πυρήνα της κυτταρικής λειτουργίας, ο χαρακτηρισμός των οποίων συντελεί στην δημιουργία ενός πλαισίου για την μοριακή βιολογία συστημάτων. Ωστόσο, γνωστές και προβλεπόμενες αλληλεπιδράσεις βρίσκονται διάσπαρτα τοποθετημένες σε διαφορετικές πηγές ενώ τα δεδομένα αυτά εμφανίζουν αξιοσημείωτες διαφορές σε θέματα ποιότητας και πληρότητας. Για την σύνδεση των διαφορεικά εκφρασμένων γονιδίων που εντοπίσαμε λοιπόν, χρησιμοποιήσαμε την βάση δεδομένων STRING v 10.0 με στόχο την σύνθεση και την αξιολόγηση των πιθανών αλληλεπιδράσεων μεταξύ των εν λόγω πρωτεϊνών, συμπεριλαμβανομένων άμεσων (φυσικών) και έμμεσων (λειτουργικών) συσχετισμών (Szkarczyk et al., 2015). 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 (Πίνακας 28). Τα διαφορεικά εκφραζόμενα γονίδια αναπαρίστανται ως κόμβοι (*nodes*) και οι διασυνδέσεις μεταξύ τους ως ακμές στο δίκτυο (Εικόνα 37,).

Πίνακας 28: Αριθμός γονιδίων που αντιστοιχήθηκαν στην βάση δεδομένων.

Group	Control_High	Control_Low
Genes	236	179

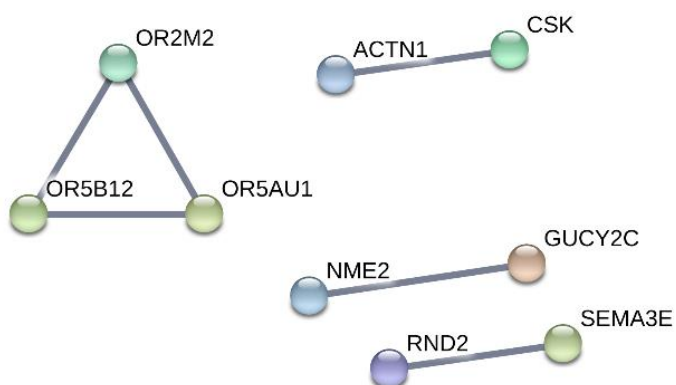
Σχετικά με τις πηγές της πληροφορίας για τις αλληλεπιδράσεις δεν λήφθηκαν υπόψιν μέθοδοι εξόρυξης κειμένων από την βιβλιογραφία, ομάδες γονιδίων που συχνά παρατηρούνται στις γενομικές περιοχές η μία της άλλης αλλά και οικογένειες γονιδίων των οποίων τα μοτίβα εμφάνισης στο γονιδίωμα εμφανίζουν ομοιότητες. 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 (Πίνακας 29) υποδηλώνουν την σημαντικότητα των ευρημάτων. Από την ανάλυση εμπλουτισμού μονοπατιών (*pathway enrichment*) βρέθηκε μόνο για την ομάδα έκθεσης σε χαμηλή δόση ένα αποτέλεσμα (Πίνακας 30).

Πίνακας 29: Στατιστικά στοιχεία για την απεικόνιση των βιολογικών δικτύων από το STRING.

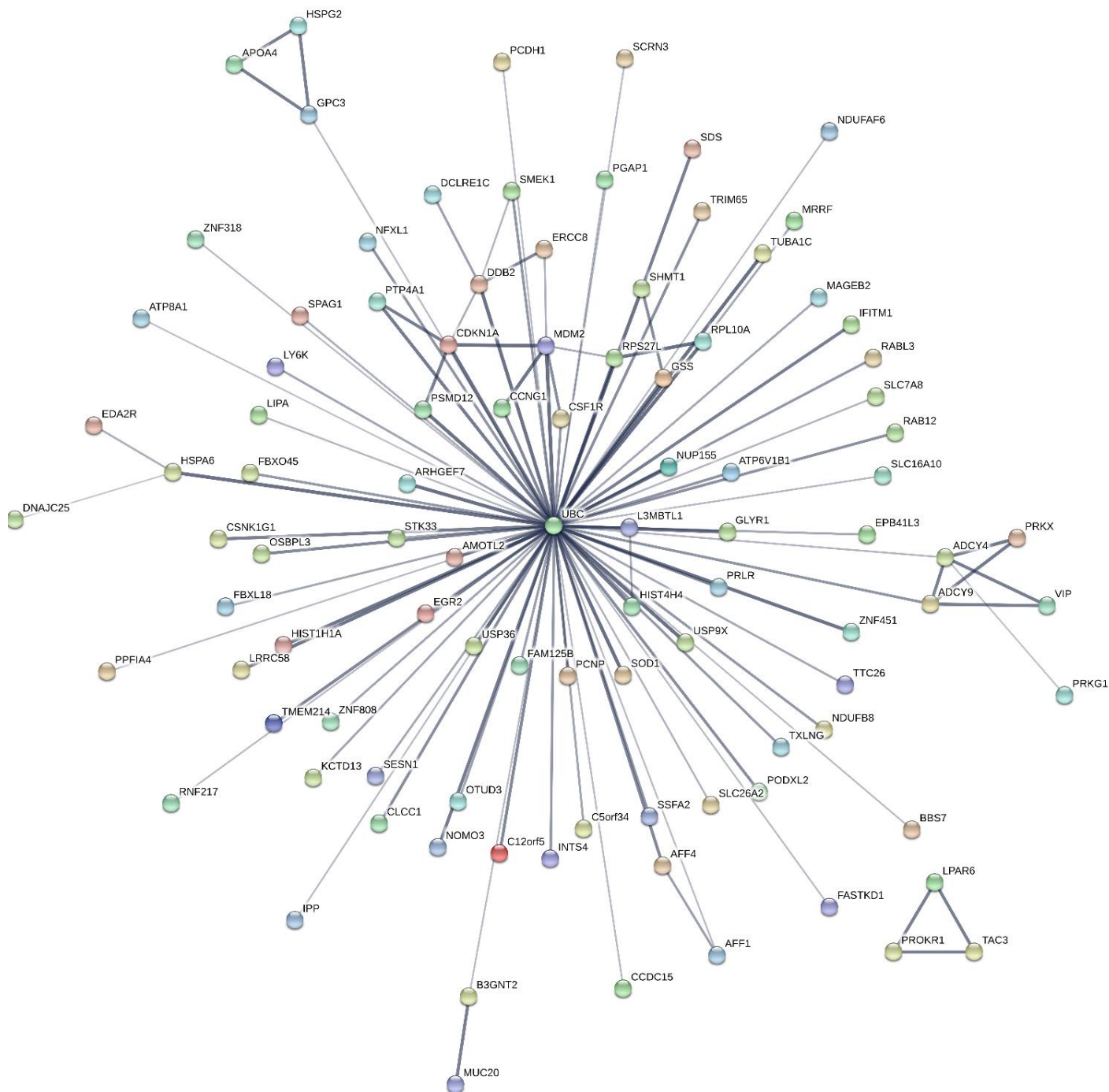
	Control_High	Control_Low
Nodes	236	179
Edges	68	6
Aver. Node Degree	0.576	0.067
Clustering Coefficient	0.98	1
Expected Edges	56	12
PPI Enrichment P_value	0.0598	0.98

Πίνακας 30: Pathway enrichment για το σύνολο γονιδίων που προέκυψαν από δείγματα τα οποία εκτέθηκαν σε χαμηλή δόση ακτινοβολίας (<0.5Gy).

Pathway ID	Pathway Description	Observed Gene Count	False Discovery Rate	Matching Proteins in Network (labels)
GO.0032399	HECT domain binding	3	0.00271	BPY2,BPY2B,BPY2C



Εικόνα 37: Βιολογικό δίκτυο των διαφορετικά εκφρασμένων γονιδίων όπως προέκυψαν από το σύνολο των δειγμάτων που εκτέθηκαν σε χαμηλή δόση ακτινοβολίας.



Εικόνα 38: Βιολογικό δίκτυο των διαφορικά εκφρασμένων γονιδίων όπως προέκυψαν από το σύνολο των δειγμάτων που εκτέθηκαν σε υψηλή δόση ακτινοβολίας.

Η σύζευξη μονομερών ή πολυμερών ουμπικουϊτίνης (UBC) μπορεί να οδηγήσει σε διάφορα αποτελέσματα μέσα σε ένα κύτταρο. Η δράση της UBC συσχετίζεται με την αποικοδόμηση πρωτεϊνών, την επιδιόρθωση DNA, τη ρύθμιση του κυτταρικού κύκλου, την τροποποίηση των κινασών, την ενδοκύττωση και τη ρύθμιση άλλων οδών σηματοδότησης κυττάρων, που δικαιολογεί την κεντρική της θέση στο δίκτυο.

6.5 Συνοπτική Παρουσίαση των Αποτελεσμάτων

Πίνακας 31: Σύνοψη αποτελεσμάτων για την έκθεση σε υψηλή δόση ακτινοβολίας (>0.5Gy).

Group	Pathways	Genes	Statistics	Analysis	Regulation
Control-High	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	MDM2 CDKN1A UBC TRIAP1 PSMD12	rawP=0.0003 adjP=0.0161	GO	n/a
	signal transduction involved in mitotic cell cycle G1/S transition DNA damage checkpoint	MDM2 CDKN1A UBC TRIAP1 PSMD12	rawP=0.0003 adjP=0.0161	GO	n/a
	L-serine catabolic process	SHMT1 SDS	rawP=0.0002 adjP=0.0161	GO	n/a
	DNA damage response (only ATM dependent)	CDKN1A FRAT1 LDLR	rawP=0.0488 adjP=0.0488	Wiki	n/a
	Gap junction	PRKX TUBA1C ADCY4 ADCY9 PRKG1	rawP=0.0016 adjP=0.0752	KEGG	n/a
	p53 signaling pathway	CCNG1 MDM2 CDKN1A SESN1	rawP=0.0041 adjP=0.0964	KEGG	n/a
	p53 pathway	GDF15 CCNG1 MDM2 CDKN1A RPS27L C12orf5 SESN1 CSNK1G1 TRIAP1	rawP=8.73e-05 adjP=0.0091	PC	n/a
	Direct p53 effectors	GDF15 CCNG1 MDM2 CDKN1A RPS27L C12orf5 SESN1 TRIAP1	rawP=5.50e-05 adjP=0.0091	PC	n/a
	Signaling by EGFR	MDM2 CDKN1A ADCY4 ARHGEF7 UBC ADCY9	rawP=0.0004 adjP=0.0279	PC	n/a
	ATR signaling pathway	GDF15 CCNG1 MDM2 CDKN1A RPS27L C12orf5 SESN1 CSNK1G1 TRIAP1	rawP=0.0007 adjP=0.0366	PC	n/a
	Target	Genes	Statistics	Analysis	Regulation
	PAX3_B	PCNP MRRF AFF4 CDKN1A EGR2 GLYR1	rawP=0.0068 adjP=0.0068	TF	↓
	FREAC3_01	CDKN1A RUNX1T1 SLC35A2 ICAM5 SCRIN3 EGR2 MAML3	rawP=0.0249 adjP=0.0249	TF	↓
	MIR-142-3P	AFF1 PCGF3 PRLR EGR2 ADCY9	rawP=0.0174 adjP=0.0174	MicroRNA	↓
	MIR-432	PCGF3 SLC7A8 PRKG1	rawP=0.0276 adjP=0.0276	MicroRNA	↓
	MIR-485-3P	AFF4 PCGF3 GLS2 RUNX1T1 MAML3	rawP=0.0320 adjP=0.0320	MicroRNA	↓

Πίνακας 32: Σύνοψη αποτελεσμάτων για την έκθεση σε χαμηλή δόση ακτινοβολίας (<0.5Gy).

Group	Pathways	Genes	Statistics	Analysis	Regulation
Control-Low	Fascia Adherens	ACTN1 NRAP	rawP=0.0059 adjP=0.0915	GO	n/a
	Intercalated Disc	ANXA5 ACTN1 NRAP	rawP=0.0032 adjP=0.0915	GO	n/a
	Cell-Cell Contact Zone	ANXA5 ACTN1 NRAP	rawP=0.0038 adjP=0.0915	GO	n/a
	Recycling Endosome	SLC9A9 VIPAS39 ZFYVE27 LAMP5 LDLR	rawP=0.0004 adjP=0.0620	GO	n/a
	Growth Cone Membrane	ZFYVE27 LAMP5	rawP=0.0022 adjP=0.0915	GO	n/a
	Pole Plasm	TRD1 TRD7	rawP=0.0059 adjP=0.0915	GO	n/a
	Germ Plasm	TRD1 TRD7	rawP=0.0059 adjP=0.0915	GO	n/a
	P Granule	TRD1 TRD7	rawP=0.0059 adjP=0.0915	GO	n/a
	Chromatoid Body	TRD1 TRD7	rawP=0.0028 adjP=0.0915	GO	n/a
	Focal Adhesion Assembly	ACTN1 SORBS1 LAMA5	rawP=0.0069 adjP=0.0069	GO	n/a
	protein localization to membrane	ZFYVE27 SORBS1 LAMA5 NRXN2	rawP=0.0109 adjP=0.0109	GO	n/a
	cell-substrate junction assembly	ACTN1 SORBS1 LAMA5	rawP=0.0159 adjP=0.0159	GO	n/a
	embryonic skeletal system development	SULF1 HOXC9 OSR2 HOXA2	rawP=0.0165 adjP=0.0165	GO	n/a
	middle ear morphogenesis	OSR2 HOXA2	rawP=0.0175 adjP=0.0175	GO	n/a
	cell junction assembly	CNTNAP1 ACTN1 SORBS1 LAMA5 SNAI2	rawP=0.0197 adjP=0.0197	GO	n/a
	NAD metabolic process	NADSYN1 DLST	rawP=0.0240 adjP=0.0240	GO	n/a
	cellular response to radiation	CDKN1A CRY1 SNAI2	rawP=0.0253 adjP=0.0253	GO	n/a
	cellular response to ionizing radiation	CDKN1A SNAI2	rawP=0.0374 adjP=0.0374	GO	n/a
	gene silencing	TDRD1 UBR2 LIN28A	rawP=0.0407 adjP=0.0407	GO	n/a
	vinculin binding	ACTN1 NRAP	rawP=0.0040 adjP=0.0040	GO	n/a
	cyclin binding	GAK CDKN1A	rawP=0.0078 adjP=0.0078	GO	n/a
	proline-rich region binding	CSK BAIAP2	rawP=0.0089 adjP=0.0089	GO	n/a
	heat shock protein binding	GAK NASP DLST TOMM34	rawP=0.0108 adjP=0.0108	GO	n/a
	oxidoreductase activity, acting on the CH-CH group of donors, NAD or NADP as acceptor	DHDH BLVRB	rawP=0.0188 adjP=0.0188	GO	n/a
	lipoprotein particle binding	APOL2 LDLR	rawP=0.0188 adjP=0.0188	GO	n/a
	protein-lipid complex binding	APOL2	rawP=0.0188 adjP=0.0188	GO	n/a
	DNA damage response (only ATM dependent)	CDKN1A FRAT1 LDLR	rawP=0.0498 adjP=0.0498	Wiki	n/a
	GAB1 signalosome	CDKN1A CSK	rawP=0.0353 adjP=0.0353	PC	n/a

Adherens junction	ACTN1 SORBS1 BAIAP2 SNAI2 FARP2	rawP=0.0005 adjP=0.0070	KEGG	n/a
HECT Domain Binding	BPY2 BPY2B BPY2C	FDR 0.00271	GO	n/a
Target	Genes	Statistics	Analysis	Regulation
ROAZ_01	BEX1 CDKN1A	rawP=0.0011 adjP=0.0011	TF	↓
TCF11MAFG_01	POPDC3 LRRFIP2 NRXN2 BLVRB	rawP=0.0181 adjP=0.0181	TF	↓
DR3_Q4	MAGEH1 ARMCX6 FRMD5 LIN28A	rawP=0.0026 adjP=0.0026	TF	↑
SOX5_01	ARMCX6 GRHL2 NOVA1 HOXA2 RBFOX2	rawP=0.0047 adjP=0.0047	TF	↑
MEF2_04	CYP46A1 NRAP	rawP=0.0052 adjP=0.0052	TF	↑
UNKNOWN	KRT222 OSR2 LIN28A	rawP=0.0053 adjP=0.0053	TF	↑
CHX10_01	KRT222 STK3 HOXA2 RBFOX2 LIN28A SULF1 OSR2 TMTC2 NOVA1	rawP=0.0055 adjP=0.0055	TF	↑
LXR_DR4_Q3	SULF1 FRMD5 LIN28A	rawP=0.0063 adjP=0.0063	TF	↑
MIR-189	KPNA4 RBM33	rawP=0.0085 adjP=0.0085	MicroRNA	↓
MIR-518A-2	EIF4G3 KPNA4 RBM33 RAB22A	rawP=0.0184 adjP=0.0184	MicroRNA	↓
MIR-181A MIR-181B MIR-181C MIR-181D	FAM122B ZNF468 CSNK1G1 OSBPL3 ZNF655 BAIAP2	rawP=0.0294 adjP=0.0294	MicroRNA	↓
MIR-518A-2	GUCY2C SORBS1 CDK2AP1 NOVA1	rawP=0.0115 adjP=0.0115	MicroRNA	↑
MIR-18A MIR-18B	GRHL2 LIN28A SH3BP4	rawP=0.0134 adjP=0.0134	MicroRNA	↑

Τα αποτελέσματά μας δείχνουν ότι έπειτα από έκθεση σε υψηλή δόση ακτινοβολίας στο πρώτο διάστημα ωρών φαίνεται να ενεργοποιούνται μονοπάτια που σχετίζονται με την ογκο-κατασταλτική πρωτεΐνη p53 (Πίνακας 12, Πίνακας 13, Πίνακας 14). Εμφανίζεται να συμβαίνει μία διαδικασία μεταγωγής σημάτων που εμπλέκονται στη μετάβαση G1/S του μιτωτικού κυτταρικού κύκλου, που αποτελεί σημείο ελέγχου βλαβών DNA (Εικόνα 30) ενώ είναι ξεκάθαρη η ενεργοποίηση μονοπατιών επιδιόρθωσης DNA (Πίνακας 12, Πίνακας 14, Πίνακας 15). Επιπλέον φαίνεται να ενεργοποιούνται χημικές αντιδράσεις και μονοπάτια που οδηγούν στον σχηματισμό της L-σερίνης. Τέλος ενεργοποιούνται μονοπάτια που σχετίζονται με τους χασμοσυνδέσμους (*gap junctions*) (Πίνακας 13, Εικόνα 31).

Η επίδραση των χαμηλών δόσεων δεν είναι τόσο ξεκάθαρη. Αρχικά σε σχέση με τα δομικά μέρη του κυττάρου (Πίνακας 18) φαίνεται να επηρεάζεται ο «*fascia adherens*», που αποτελεί ένα σύνδεσμο κυττάρου-κυττάρου που περιέχει την διαμεμβρανική πρωτεΐνη N-καδερίνη, η οποία αλληλεπιδρά με πανομοιότυπα μόρια από γειτονικά

κύτταρα και σχηματίζει έναν ισχυρό μηχανικό διακυτταρικό δεσμό με αποτέλεσμα τη δημιουργία μιας δομής μυοϊνιδίων (*myofibrils*) που καταλήγουν σε καρδιομυοκύτταρα. Εμφανίζονται επιδράσεις στο ενδόσωμα ανακύκλωσης, ένα οργανίδιο που αποτελείται από ένα δίκτυο σωλήνων που λειτουργεί στοχεύοντας μόρια, όπως υποδοχείς, μεταφορείς και λιπίδια, στην κυτταροπλασματική μεμβράνη και την αντίστοιχη λιπιδική διπλοστοιβάδα που περιβάλλει ένα ενδόσωμα ανακύκλωσης. Επίσης αλλαγές μπορεί να σημειωθούν στο τμήμα της κυτταροπλασματικής μεμβράνης που περιβάλλει έναν κώνο ανάπτυξης (*growth cone*), μεγάλη επέκταση που υποστηρίζεται από ακτίνη, ενός αναπτυσσόμενου ή αναγεννώμενου νευρίτη που αναζητεί τον στόχο σύναψής του. Πιθανό είναι να επηρεάζονται και διαδικασίες κυτταρικής διαφοροποίησης κατά την ανάπτυξη του εμβρύου σε περίπτωση εγκυμοσύνης .

Στο κομμάτι των μοριακών λειτουργιών (Πίνακας 20) εμφανίζονται μεταβολές σε πολλαπλές επιλεκτικές και μη ομοιοπολικές αλληλεπιδράσεις. Τέτοιες είναι αρχικά η αλληλεπίδραση με την βινκουλίνη, μια πρωτεΐνη που βρίσκεται στους μυς, στους ινοβλάστες και τα επιθηλιακά κύτταρα, η οποία δεσμεύει την ακτίνη και φαίνεται να μεσολαβεί στη σύνδεση των ινών ακτίνης με πρωτεΐνες της μεμβράνης του πλάσματος. Επιπλέον εμφανίζεται η αλληλεπίδραση με τις κυκλίνες, πρωτεΐνες των οποίων τα επίπεδα σε ένα κύτταρο μεταβάλλονται σημαντικά κατά τη διάρκεια του κυτταρικού κύκλου, αυξάνονται σταθερά μέχρι τη μίτωση και στη συνέχεια πέφτουν απότομα στο μηδέν. Καθώς οι κυκλίνες φθάνουν σε ένα επίπεδο κατωφλίου, πιστεύεται ότι οδηγούν τα κύτταρα στη φάση G2 και έτσι στη μίτωση. Άλλες αλληλεπιδράσεις είναι με περιοχές πλούσιες σε προλίνη, με σύμπλοκα πρωτεΐνης-λιπιδίου, με σωματίδια λιποπρωτεΐνης ή με πρωτεΐνες θερμικού σοκ, που συντίθενται ή ενεργοποιούνται, ως απόκριση σε θερμικό σοκ. Τέλος επηρεάζονται οι αντιδράσεις οξειδοαναγωγής κατά τις οποίες μια ομάδα CH-CH δρα ως δότης υδρογόνου ή ηλεκτρονίου και μειώνει το NAD ή το NADP. Το NAD είναι ένα δινουκλεοτίδιο που αποτελείται από δύο νουκλεοτίδια ενωμένα μέσω των φωσφορικών τους ομάδων. Το ένα νουκλεοτίδιο περιέχει μια βάση αδενίνης και το άλλο ένα νικοτιναμίδιο, υπάρχει σε δύο μορφές, μία οξειδωμένη και μία μειωμένη μορφή NAD⁺ και NADH αντίστοιχα. Το NADP περιέχει μία επιπλέον φωσφορική ομάδα.

Παρόλα αυτά στο κομμάτι των βιολογικών διεργασιών (Πίνακας 19) προκύπτουν ορισμένα πιο ξεκάθαρα και αναμενόμενα αποτελέσματα. Αρχικά φαίνεται να επηρεάζονται οι διαδικασίες συσσωμάτωσης και συγκόλλησης συστατικών που έχουν ως στόχο να σχηματίσουν μια εστιακή προσκόλληση (*focal adhesion*), η οποία αποτελεί ένα σύμπλεγμα ενδοκυτταρικών σηματοδοτικών και δομικών πρωτεϊνών που

παρέχει δομική σύνδεση μεταξύ του εσωτερικού κυτταροσκελετού ακτίνης και της ECM και λειτουργεί ως τοποθεσία μεταγωγής σήματος. Κυριότερα όμως διαφαίνεται η επίδραση σε διαδικασίες αλλαγής της κατάστασης ή της δραστηριότητας ενός κυττάρου, από την άποψη της κίνησης, της έκκρισης, της παραγωγής ενζύμων, της γονιδιακής έκφρασης κ.λπ. ως αποτέλεσμα ενός ερεθίσματος ηλεκτρομαγνητικής και ιοντίζουσας ακτινοβολίας. Άλλα ενδιαφέροντα αποτελέσματα που προέκυψαν είναι η ενεργοποίηση διεργασιών που διεξάγονται σε κυτταρικό επίπεδο με αποτέλεσμα είτε τη μακροπρόθεσμη καταστολή της μεταγραφής μέσω δράσεως επί της δομής της χρωματίνης ή την καταστολή της γονιδιακής έκφρασης μετά την μεταγραφή, με μεσολάβηση του RNA (*gene silencing*) καθώς και ο μεταβολισμός του NAD.

Προκύπτει επίσης το «*DNA damage response (only ATM dependent)*», το δεύτερο από δύο μονοπάτια που σχετίζονται με την απόκριση σε βλάβες DNA (Πίνακας 21) όπως και στην περίπτωση των υψηλών δόσεων. Τέλος μεγάλο ενδιαφέρον εμφανίζει το μονοπάτι «*GAB1 signalosome*» (Πίνακας 22) καθώς το γονίδιο GAB1 διαδραματίζει κεντρικό ρόλο στην απόκριση κυτταρικής ανάπτυξης, τον μετασχηματισμό και την απόπτωση αλλά και το μονοπάτι «*HECT domain binding*» (Πίνακας 30, Εικόνα 33). Τέλος η έκθεση σε χαμηλές δόσεις φαίνεται να επηρεάζει και τους *adherens junctions* (AJs) ή συνδέσμους προσκόλλησης κυττάρων. Αυτοί αποτελούν τον πιο κοινό τύπο διακυτταρικών προσκολλήσεων-συμφύσεων, και είναι σημαντικοί για τη διατήρηση της δομής των ιστών και της πολικότητας των κυττάρων, ενώ έχουν την δυνατότητα να περιορίζουν την κίνηση και τον πολλαπλασιασμό αυτών. (Εικόνα 36).

6.6 Γονίδια από την Βιβλιογραφία

Τα γονίδια (Πίνακας 33, Πίνακας 34) από την βιβλιογραφική αναζήτηση στο Quertle, αντιστοιχήθηκαν με το αντίστοιχο HGNC και την αντίστοιχη περιγραφή από το BioMart.

Πίνακας 33: Διαφορικά εκφραζόμενα γονίδια που προκύπτουν από την βιβλιογραφία για τις υψηλές 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	FBXO22	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 34: Διαφορικά εκφραζόμενα γονίδια που προκύπτουν από την βιβλιογραφία για τις χαμηλές δόσεις (<0.5Gy).

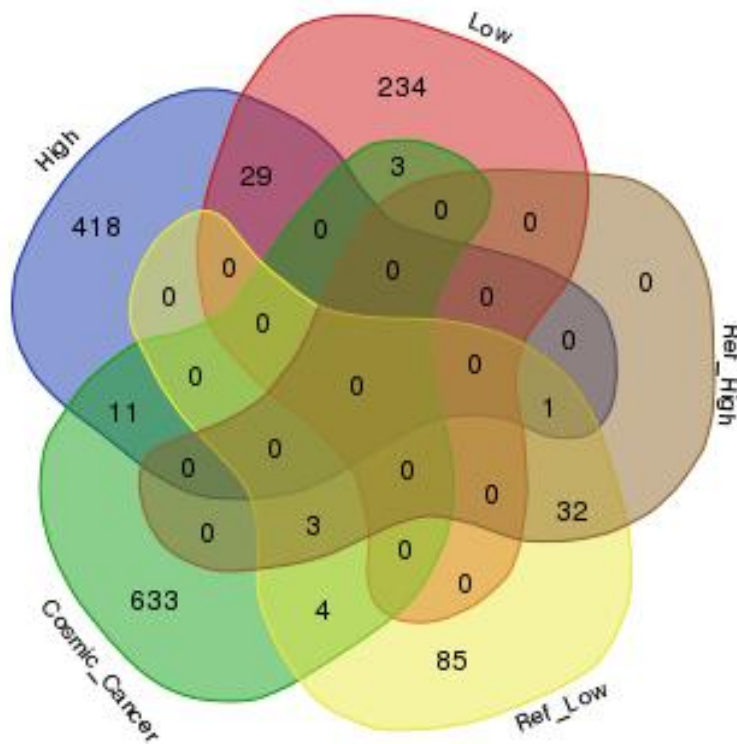
ENSG_ID	HGNC_Symbol	Description
ENSG00000163017	ACTG2	actin, gamma 2, smooth muscle, enteric [Source:HGNC Symbol;Acc:HGNC:145]
ENSG0000020181	ADGRA2	adhesion G protein-coupled receptor A2 [Source:HGNC Symbol;Acc:HGNC:17849]
ENSG00000205336	ADGRG1	adhesion G protein-coupled receptor G1 [Source:HGNC Symbol;Acc:HGNC:4512]
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	dolichyldiphosphatase 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]

ENSG0000005238	FAM214B	family with sequence similarity 214 member B [Source:HGNC Symbol;Acc:HGNC:25666]
ENSG00000167196	FBXO22	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	mucolipin 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]
ENSG00000181889	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	ZDHC6	zinc finger DHC-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]

6.7 Τελικό Διάγραμμα Venn

Η κατασκευή του διαγράμματος Venn (Εικόνα 39) δείχνει γονίδια που εκφράζονται διαφορετικά έπειτα από έκθεση σε υψηλή και χαμηλή δόση (Πίνακας 36), με όριο διαχωρισμού αυτών τα 0.5Gy, όπως προέκυψαν από την ανάλυση αυτή. Σημειώνεται πως οι μελέτες που επιλέχθηκαν κατά την βιβλιογραφική αναζήτηση επικεντρώνονταν στις χαμηλές δόσεις ακτινοβολίας. Για αυτό το λόγο τα γονίδια που ξεχωρίζουν στις χαμηλές δόσεις είναι ποσοτικά περισσότερα από αυτά στις υψηλές, γεγονός που δεν είναι συμβατό με την γενική πραγματικότητα.



Εικόνα 39: Τελικό διάγραμμα Venn όπου συνδυάζονται τα γονίδια που ξεχώρισαν σε αυτή την εργασία με τα γονίδια από την βιβλιογραφία και την λίστα γονιδίων που σχετίζονται με την δημιουργία καρκίνου. Το όριο για τις χαμηλές δόσεις είναι τα 0.5Gy.

Βρέθηκαν 6 γονίδια κοινά με την λίστα γονιδίων που σχετίζονται με τον καρκίνο στις υψηλές (Πίνακας 38) και 3 γονίδια στις χαμηλές δόσεις ακτινοβολίας (Πίνακας 39). Τέλος βρέθηκε ένα μόνο γονίδιο κοινό με την βιβλιογραφία στις υψηλές δόσεις (Πίνακας 35) και 3 γονίδια κοινά μεταξύ της βιβλιογραφίας και των γονιδίων που σχετίζονται με τον καρκίνο (Πίνακας 37).

Πίνακας 35: Γονίδια που βρέθηκαν κοινά μεταξύ της μελέτης μας για την έκθεση σε υψηλές δόσεις με την βιβλιογραφία.

High_Reference		
ENSG_ID	HGNC	Description
ENSG00000164331	ANKRA2	Ankyrin Repeat Family A Member 2 [Source:HGNC Symbol;Acc:HGNC:13208]

Πίνακας 36: Γονίδια που βρέθηκαν στην μελέτη μας κοινά για την έκθεση σε υψηλές και χαμηλές δόσεις.

High_Low		
ENSG_ID	HGNC	Description
ENSG00000256614	AK6P1	adenylate kinase 6 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:51678]
ENSG00000187997	C17orf99	chromosome 17 open reading frame 99 [Source:HGNC Symbol;Acc:HGNC:34490]
ENSG00000124762	CDKN1A	cyclin dependent kinase inhibitor 1A [Source:HGNC Symbol;Acc:HGNC:1784]
ENSG00000169118	CSNK1G1	casein kinase 1 gamma 1 [Source:HGNC Symbol;Acc:HGNC:2454]
ENSG00000236417	CTSLP1	cathepsin L pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:2539]
ENSG00000230011	CTSLP4	cathepsin L pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:23645]
ENSG00000204437	CTSLP6	cathepsin L pseudogene 6 [Source:HGNC Symbol;Acc:HGNC:23611]
ENSG00000135423	GLS2	glutaminase 2 [Source:HGNC Symbol;Acc:HGNC:29570]
ENSG00000105376	ICAM5	intercellular adhesion molecule 5 [Source:HGNC Symbol;Acc:HGNC:5348]
ENSG00000185885	IFITM1	interferon induced transmembrane protein 1 [Source:HGNC Symbol;Acc:HGNC:5412]
ENSG00000274843	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000277753	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000176945	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000275430	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000281630	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000276583	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000275501	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000278114	MUC20	mucin 20, cell surface associated [Source:HGNC Symbol;Acc:HGNC:23282]
ENSG00000070882	OSBPL3	oxysterol binding protein like 3 [Source:HGNC Symbol;Acc:HGNC:16370]
ENSG00000185619	PCGF3	polycomb group ring finger 3 [Source:HGNC Symbol;Acc:HGNC:10066]
ENSG00000237872	POU5F1P4	POU class 5 homeobox 1 pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:33310]
ENSG00000112299	VNN1	vanin 1 [Source:HGNC Symbol;Acc:HGNC:12705]
ENSG00000130684	ZNF337	zinc finger protein 337 [Source:HGNC Symbol;Acc:HGNC:15809]
ENSG00000196152	ZNF79	zinc finger protein 79 [Source:HGNC Symbol;Acc:HGNC:13153]
ENSG00000106479	ZNF862	zinc finger protein 862 [Source:HGNC Symbol;Acc:HGNC:34519]
ENSG00000235978		
ENSG00000268858		
ENSG00000239665		
ENSG00000263293		thyroid cancer-associated transcript 158 [Source:EntrezGene;Acc:102724508]

Πίνακας 37: Γονίδια που εντοπίσαμε από την βιβλιογραφία και συμπίπτουν με την λίστα για τα γονίδια που σχετίζονται με την εμφάνιση καρκίνου

Cosmic_Reference		
ENSG_ID	HGNC	Description
ENSG00000136167	LCP1	lymphocyte cytosolic protein 1 [Source:HGNC Symbol;Acc:HGNC:6528]
ENSG00000147065	MSN	moesin [Source:HGNC Symbol;Acc:HGNC:7373]
ENSG00000154767	XPC	XPC complex subunit, DNA damage recognition and repair factor [Source:HGNC Symbol;Acc:HGNC:12816]

Πίνακας 38: Γονίδια που ξεχωρίσαμε για την έκθεση σε υψηλές δόσεις και συμπίπτουν με την λίστα για τα γονίδια που σχετίζονται με την εμφάνιση καρκίνου.

Cosmic_High		
ENSG_ID	HGNC	Description
ENSG00000172493	AFF1	AF4/FMR2 family member 1 [Source:HGNC Symbol;Acc:HGNC:7135]
ENSG00000072364	AFF4	AF4/FMR2 family member 4 [Source:HGNC Symbol;Acc:HGNC:17869]
ENSG00000135679	MDM2	MDM2 proto-oncogene [Source:HGNC Symbol;Acc:HGNC:6973]
ENSG00000235068	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]
ENSG00000230336	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]
ENSG00000237582	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]
ENSG00000233911	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]
ENSG00000206454	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]
ENSG00000229094	POU5F1	POU class 5 homeobox 1 [Source:HGNC Symbol;Acc:HGNC:9221]
ENSG00000079102	RUNX1T1	RUNX1 translocation partner 1 [Source:HGNC Symbol;Acc:HGNC:1535]
ENSG00000107807	TLX1	T-cell leukemia homeobox 1 [Source:HGNC Symbol;Acc:HGNC:5056]

Πίνακας 39: Γονίδια που ξεχωρίσαμε για την έκθεση σε χαμηλές δόσεις και συμπίπτουν με την λίστα για τα γονίδια που σχετίζονται με την εμφάνιση καρκίνου.

Cosmic_Low		
ENSG_ID	HGNC	Description
ENSG00000108091	CCDC6	coiled-coil domain containing 6 [Source:HGNC Symbol;Acc:HGNC:18782]
ENSG00000169714	CNBP	CCHC-type zinc finger nucleic acid binding protein [Source:HGNC Symbol;Acc:HGNC:13164]
ENSG00000169249	ZRSR2	zinc finger CCCH-type, RNA binding motif and serine/arginine rich 2 [Source:HGNC Symbol;Acc:HGNC:23019]

7 ΣΥΖΗΤΗΣΗ

Στα αποτελέσματα που προκύπτουν για την απόκριση σε έκθεση χαμηλής δόσης ιοντίζουσας ακτινοβολίας, υπάρχουν ενδείξεις ότι ενεργοποιούνται μηχανισμοί επιδιόρθωσης DNA. Σε αυτά συμπεριλαμβάνεται κατά προφανή τρόπο το «DNA damage response (only ATM dependent)», που σχετίζεται με τους μηχανισμούς επιδιόρθωσης βλαβών DNA. Το μονοπάτι αυτό διαθέτει δύο κεντρικά γονιδιακά προϊόντα (την ATM και την p53), που συνδέονται με το πρώτο μονοπάτι απόκρισης βλαβών DNA. Ο στόχος αυτού είναι η αναφορά περισσότερων γονιδιακών προϊόντων και αλλαγών στην κατάσταση των κυττάρων μέσω της οδού απόκρισης σε βλάβες DNA και η ταυτόχρονη διατήρηση της αλληλουχίας του DNA (Kitagawa and Kastan, 2005). Αναμενόμενη είναι επίσης η ενεργοποίηση μονοπατιών απόκρισης σε ερέθισμα ηλεκτρομαγνητικής και ιοντίζουσας ακτινοβολίας, ενώ το σταμάτημα της έκφρασης γονιδίων αποτελεί επίσης έναν πιθανό τρόπο άμυνας του κυττάρου.

Τα συμπεράσματα για τις κυτταρικές δομές που προκύπτουν δεν είναι ξεκάθαρα. Από τις διαδικασίες που σχετίζονται με τις μοριακές λειτουργίες του κυττάρου, έμφαση πρέπει να δωθεί στις διαδικασίες που σηματοδοτούν την αλλαγή φάσης στον κυτταρικό κύκλο, δεδομένου ότι μπορούν και αυτές με την σειρά τους να επηρεάσουν τις διαδικασίες αντιγραφής, μεταγραφής και μετάφρασης που συμβαίνουν κατά την κυτταρική διαίρεση και έτσι μπορούν να καθορίσουν την επιτυχή επιδιόρθωση βλαβών και κατά συνέπεια την διαιώνιση ή μη, κάποιας λανθασμένης αλληλουχίας. Η απόκριση σε θερμικό σοκ είναι και αυτή αναμενόμενη, καθώς κατά την αλληλεπίδραση της ακτινοβολίας με την έμβια ύλη σημειώνεται αύξηση της θερμοκρασίας. Τέλος οι οξειδοαναγωγικές αντιδράσεις με υποδοχέα το NAD αποτελούν ένδειξη οξειδωτικού στρες. Οι μεταβολές στις βιολογικές διεργασίες που εμφανίζονται, σχετίζονται με την μεταγωγή σημάτων. Ειδικά οι σύνδεσμοι προσκόλλησης κυττάρων δίνουν βάση για περαιτέρω μελέτη γύρω από το φαινόμενο του παρατυχόντος (*bystander effect*). Φαίνεται ακόμα να ενεργοποιούνται διαδικασίες περιορισμού του κυτταρικού πολλαπλασιασμού, που μπορεί να αποτελούν κομμάτι επιδιορθωτικών μηχανισμών.

Το μονοπάτι «GAB1 signalosome» παρουσιάζει μεγάλο ενδιαφέρον, καθώς το GAB1 δρα ως πρωτεΐνη προσαρμογής που σχετίζεται με την ανάπτυξη του κυττάρου, τον μετασχηματισμό και την απόπτωση. Οδηγεί στον σχηματισμό ενός ενεργού PIK3, το οποίο σε σύμπλοκο με τα EGFR, GRB2 και GAB1, καταλύει τη φωσφορυλίωση της PIP2 και τη μετατροπή της σε PIP3, γεγονός που οδηγεί στην ενεργοποίηση της σηματοδότησης της AKT. Η κινάση σερίνης/θρεονίνης AKT (γνωστή και ως πρωτεϊνική κινάση B ή PKB) από την αρχική ανακάλυψή της ως πρωτο-ογκογονίδιο, έχει καταστεί

κύριο επίκεντρο προσοχής λόγω του κρίσιμου ρόλου της στη ρύθμιση των ποικίλων κυτταρικών λειτουργιών συμπεριλαμβανομένου του μεταβολισμού, της ανάπτυξης, της επιβίωσης, της μεταγραφής και της σύνθεσης πρωτεϊνών. Επίσης θεωρείται ότι σχετίζεται με την δημιουργία καρκίνου.

Η επιλεκτική και μη-ομοιοπολική αλληλεπίδραση με μία HECT, «ομόλογη προς το καρβοξυλικό τερματικό E6-AP», περιοχή μιας πρωτεΐνης, σχετίζεται με την ρύθμιση της ουμπικουτινώσεως του CDC25. Η λιγάση ουμπικουιτίνης δέχεται ουμπικουιτίνη από ένα ένζυμο συζεύξεως ουμπικουιτίνης E2 με τη μορφή θειοεστέρα και κατόπιν μεταφέρει απευθείας την ουμπικουιτίνη σε στοχευμένα υποστρώματα. Ένα υπόλειμμα κυστεΐνης απαιτείται για τον σχηματισμό ουμπικουιτίνης-θειοεστέρος. Η πρωτεΐνη αλληλεπίδρασης του υποδοχέα ανθρώπινου θυρεοειδούς 12, η οποία περιέχει αυτή την περιοχή, αποτελεί συστατικό μιας εξαρτώμενης από ATP πρωτεΐνης πολλαπλής υπομονάδας που αλληλεπιδρά με την περιοχή δέσμευσης συνδέτη του υποδοχέα θυρεοειδούς ορμόνης. Επιπλέον η ανθρώπινη λιγάση ουμπικουιτίνης E3A αλληλεπιδρά με την πρωτεΐνη E6, η οποία συσχετίζεται με τον τύπο καρκίνου ανθρώπινου ιού θηλώματος τύπου 16 και 18. Τέλος το σύμπλοκο E6/E6-AP δεσμεύεται και κατευθύνεται προς την ογκο-κατασταλτική πρωτεΐνη p53 με σκοπό την πρωτεόλυση με μεσολάβηση ουβικουιτίνης. Οι παραπάνω περιπτώσεις ανήκουν στα σημαντικά και ισχυρά αποτελέσματα της μελέτης αυτής

Τα αποτελέσματα που προκύπτουν για τις υψηλές δόσεις θεωρούνται αναμενόμενα για την περιγραφή της απόκρισης του κυττάρου έπειτα από έκθεση σε ακτινοβολία. Η ενεργοποίηση της p53, που επάγεται από έναν αριθμό σημάτων στρες όπως οι βλάβες DNA, το οξειδωτικό στρες και τα ενεργοποιημένα ογκογονίδια είναι το βασικότερο από αυτά. Η πρωτεΐνη p53 χρησιμοποιείται ως μεταγραφικός ενεργοποιητής των γονιδίων που ρυθμίζονται από αυτήν (p53-regulated genes). Αυτό έχει τρία βασικά αποτελέσματα: την διακοπή του κυτταρικού κύκλου, την γήρανση ή την απόπτωση του κυττάρου. Μέσω άλλων λειτουργιών των γονιδίων αυτών, που σχετίζονται με την επικοινωνία με τα γειτονικά κύτταρα, επιδιορθώνονται βλάβες του DNA ή δημιουργούνται θετικοί και αρνητικοί βρόχοι ανάδρασης που ενισχύουν ή εξασθενούν τις λειτουργίες της πρωτεΐνης p53 και ενσωματώνουν αυτές τις αποκρίσεις στρες με άλλα μονοπάτια μεταγωγής σήματος (Harris and Levine, 2005). Επίσης έχουμε αναστολή του κυτταρικού κύκλου μέσω ενεργοποίησης της p53. Συμβαίνει μία αλληλουχία διεργασιών, που επάγονται από τον ρυθμιστή του κυτταρικού κύκλου την φωσφοπρωτεΐνη p53, ή άλλη ισοδύναμη πρωτεΐνη, ως απόκριση στον εντοπισμό βλάβης DNA και συντελούν στην διακοπή της διαδικασίας της μείωσης κατά την πορεία

του κυτταρικού κύκλου. Προκύπτει επίσης και το μονοπάτι απόκρισης σε βλάβες «DNA damage response (only ATM dependent)» όπως και στην περίπτωση της έκθεσης σε χαμηλή δόση ακτινοβολίας, ενώ η σηματοδότηση μέσω της κινάσης ATR αποτελεί επίσης μέρος της ενεργοποίησης των μηχανισμών επιδιόρθωσης.

Η διεργασία καταβολισμού της L-σερίνης (Πίνακας 12) όπως προκύπτει από τους όρους υπερεκπροσώπησης, διαθέτει κρίσιμο ρόλο στην εξασφάλιση της σωστής λειτουργίας του κεντρικού νευρικού συστήματος και του εγκεφάλου. Εμπλέκεται στο σχηματισμό φωσφολιπιδίων που απαιτούνται για την παραγωγή κυττάρων και διαδραματίζει σημαντικό ρόλο στη λειτουργία DNA και RNA, στο σχηματισμό των μυών και στο μεταβολισμό λιπών και λιπαρών οξέων. Κυριότερα τα αντισώματα και οι ανοσοσφαιρίνες που απαιτούνται για την υποστήριξη του ανοσοποιητικού συστήματος, που ενεργοποιείται έπειτα από έκθεση σε ιοντίζουσα ακτινοβολία, απαιτούν επίσης σερίνη για να παραχθούν (Ramos and Wiame, 1982). Συνεπώς το μονοπάτι αυτό αποτελεί ακόμα μια ένδειξη της απόκρισης σε έκθεση υψηλής δόσης ιοντίζουσας ακτινοβολίας.

Οφείλει επίσης να σημειωθεί ότι η υπερέκφραση της EGFR (*Epidermal Growth Factor Receptor*) έχει συσχετιστεί με έναν αριθμό καρκίνων, συμπεριλαμβανομένου του καρκίνου του πνεύμονα και του μαστού, του γλοιοβλαστώματος και των επιθηλιακών όγκων της κεφαλής και του αυχένα. Οι σωματικές μεταλλάξεις που εμπλέκουν την πρωτεΐνη EGFR οδηγούν στη σταθερή ενεργοποίησή του, η οποία προκαλεί ανεξέλεγκτη κυτταρική διαίρεση (Alanazi and Khan, 2016), ενώ μεταλλάξεις, ενισχύσεις και λανθασμένες ρυθμίσεις της EGFR εμπλέκονται σε περίπου 30% όλων των επιθηλιακών καρκίνων.

Οι χασμοσύνδεσμοι (*gap junctions*) πρόκεινται για διακυτταρικά κανάλια που επιτρέπουν την άμεση επικοινωνία μεταξύ των γειτονικών κυττάρων. Πιο συγκεκριμένα επιτρέπουν την άμεση μεταφορά μικρών μορίων συμπεριλαμβανομένων ιόντων, αμινοξέων, νουκλεοτιδίων, δευτερευόντων αγγελιαφόρων και άλλων μεταβολιτών. Η επικοινωνία μεταξύ των χασμοσυνδέσμων είναι σημαντική για πολλές φυσιολογικές διεργασίες, όπως η εμβρυϊκή ανάπτυξη, η ηλεκτρική σύζευξη, η μεταβολική μεταφορά, η απόπτωση, και η ομοιοστάση των ιστών. Η επικοινωνία δια μέσω αυτών επηρεάζεται από ποικιλία ερεθισμάτων, συμπεριλαμβανομένων των μεταβολών στο επίπεδο του ενδοκυτταρικού Ca_2^+ του pH, της εφαρμοζόμενης στο σημείο τάσης όπως και πολλαπλών διαδικασιών φωσφορυλίωσης ή αποφωσφορυλίωσης (Εικόνα 31). Η ενεργοποίηση αυτών σχετίζεται άμεσα με το φαινόμενο του παρατυχόντος.

Συγκρίνοντας τις επιδράσεις υψηλών και χαμηλών δόσεων ακτινοβολίας, φαίνεται πως και στις δύο περιπτώσεις ενεργοποιούνται μηχανισμοί επιδιόρθωσης. Παρόλα αυτά τα αποτελέσματα για τις υψηλές δόσεις δείχνουν ότι το κύτταρο έχει αντιληφθεί την δημιουργία βλαβών και δρα άμεσα για την επιδιόρθωση αυτών. Επιπλέον ο κίνδυνος εμφάνισης κάποιου τύπου καρκίνου είναι αυξημένος καθώς οι βλάβες από υψηλή δόση ακτινοβολίας είναι μεγάλες ποιοτικά και ποσοτικά και δεν είναι πάντοτε δυνατό να επιδιορθωθούν πλήρως ακόμα και αν το κύτταρο έχει αντιληφθεί την παρουσία τους. Παράλληλα στην περίπτωση που δεν ρυθμιστεί η έκφραση της EGFR, η πιθανότητα εμφάνισης κάποιου τύπου καρκίνου είναι υψηλή. Αντίθετα στις χαμηλές δόσεις δεν είναι τυχαίο που τα αποτελέσματα, αν και πολλαπλά, δεν δείχνουν κατά μεγάλο ποσοστό την ενεργοποίηση του κυττάρου ως προς την επιδιόρθωση. Φαίνεται παρόλα αυτά όμως, πως με μεταβολές, ενισχύσεις και γενικότερα λανθασμένες ρυθμίσεις των GAB1 και AKT όπως και των πρωτεϊνών που περιέχουν την HECT μπορεί να οδηγηθούμε σε μεταλλάξεις ή καρκινογένεση.

Παρατηρείται ότι το γονίδιο CDKN1A εμφανίζεται στην πλειοψηφία των αποτελεσμάτων μας για έκθεση σε χαμηλή δόση ακτινοβολίας. Η έκφραση του γονιδίου αυτού οδηγεί στην παραγωγή της πρωτεΐνης p21, η οποία επάγεται από την ογκο-κατασταλτική πρωτεΐνη p53. Το γεγονός αυτό δίνει ένδειξη για την ενεργοποίηση της p53, ακόμα και αν δεν αναδείχθηκε από τους όρους υπερεκπροσώπησης. Η p21 αποτελεί αναστολές κυκλινο-εξαρτώμενων κινασών, οι οποίες ευθύνονται για την πρόοδο του κυτταρικού κύκλου (Georgakilas et al., 2017). Συνεπώς η ενεργοποίηση της πρωτεΐνης p21 οδηγεί στην αναστολή του κυτταρικού κύκλου και μάλιστα συγκεκριμένα στην μετάβαση G1/S πριν την διαδικασία της αντιγραφής, με σκοπό την επιδιόρθωση πιθανών βλαβών.

Από το τελικό διάγραμμα Venn βλέπουμε ότι δεν βρέθηκε στατιστικά σημαντική υπερεκπροσώπηση. Παρόλα αυτά το γονίδιο ANKRA2 βρέθηκε κοινό με την βιβλιογραφία για τις υψηλές δόσεις, κοντινότερο παράλογο του οποίου αποτελεί το RFXANK το οποίο και αντικαθιστά κατά την ενεργοποίηση των MHC-II γονιδίων που σχετίζονται με την ανοσοαπόκριση. Συγκρίνοντας τα αποτελέσματά μας με την λίστα γονιδίων που σχετίζονται με τον καρκίνο προέκυψαν για τις υψηλές δόσεις ακτινοβολίας τα: AFF1, AFF4, MDM2, POU5F1, RUNX1T1 και TLX1. Η MDM2 μπορεί να προάγει τον σχηματισμό όγκου στοχεύοντας πρωτεΐνες καταστολής όγκων, όπως η p53. Η μεταγραφή του ίδιου του γονιδίου ρυθμίζεται από την p53 και η υπερέκφραση αυτού ανιχνεύεται σε ποικιλία τύπων καρκίνου. Το POU5F1 με την σειρά του διαδραματίζει βασικό ρόλο στην εμβρυϊκή ανάπτυξη ενώ η ανώμαλη έκφραση αυτού

του γονιδίου σε ενήλικους ιστούς συνδέεται με την ογκογένεση. Αυτό το γονίδιο μπορεί να συμμετέχει σε μετατόπιση (*translocation*) με το γονίδιο σαρκώματος του Ewing (*Ewing's Sarcoma*) στο χρωμόσωμα 21, το οποίο επίσης οδηγεί σε σχηματισμό όγκου. Τα AFF1, AFF4, RUNX1T1 και TLX1 φαίνεται να σχετίζονται άμεσα με την νόσο της λευχαιμίας. Η σύγκριση με τα αποτελέσματά μας για τις χαμηλές δόσεις έδωσε τα: CCDC6, CNBP, και ZRSR2. Η κωδικοποιούμενη πρωτεΐνη από το CCDC6 εκφράζεται παντού και μπορεί να λειτουργήσει ως ογκο-καταστολέας. Μια χρωμοσωμική αναδιάταξη η οποία καταλήγει στην έκφραση ενός γονιδίου που περιέχει ένα τμήμα αυτού του γονιδίου και την περιοχή κωδικοποίησης της ενδοκυτταρικής κινάσης του πρωτο-ογκογονιδίου του αμφιβληστροειδούς είναι η αιτία καρκινώματος στο θηλώδες του θυρεοειδούς (*thyroid papillary carcinoma*). Τέλος, το CNBP σχετίζεται με την μυοτονική δυστροφία τύπου 2 (DM2), η οποία αποτελεί μία νόσο που χαρακτηρίζεται από μυοτονία (ακούσια μυϊκή συστολή με καθυστερημένη χαλάρωση), μυϊκή δυσλειτουργία (αδυναμία, πόνο και δυσκαμψία) και λιγότερο συχνά από καρδιακή δυσλειτουργία, καταρράκτη, ανεπάρκεια όρχεων και σακχαρώδη διαβήτη τύπου 2, μη ινσουλινο-εξαρτώμενο.

Από τα αρχικά διαγράμματα Venn με χρήση των p-values που υπολογίστηκαν με Μπεϋζιανή ανάλυση, διαπιστώνεται ότι η διεξαγωγή ενός τέτοιου εγχειρήματος για την ανάδειξη διαφορικά εκφρασμένων γονιδίων έπειτα από έκθεση σε ιονίζουσα ακτινοβολία χρήζει αναθεώρησης. Το γεγονός ότι δεν εμφανίζονται στατιστικώς σημαντικά διαφορικώς εκφραζόμενα γονίδια σε όλες τις μελέτες μέσω Μπεϋζιανής ανάλυσης δείχνει τον πιθανό λανθασμένο πειραματικό σχεδιασμό των μελετών ενώ αποτελεί μία εξήγηση για την αποτυχία των μεθόδων διόρθωσης πολλαπλών ερωτημάτων κατά την εφαρμογή μετα-ανάλυσης.

Η αποδοχή των γονιδίων που προέκυψαν από την μέθοδο της μετα-ανάλυσης ως στατιστικά σημαντικών συνεπώς δεν είναι απολύτως σωστή, καθώς δεν επιβεβαιώθηκαν από καμία μέθοδο διόρθωσης πολλαπλών ερωτημάτων. Θεωρούμε ότι το πρόβλημα δεν εντοπίζεται στη μέθοδο αυτής της ανάλυσης, καθώς τα δεδομένα μας λήφθηκαν από προϋπάρχοντα πειράματα που χρησιμοποιούσαν διαφορετικές τεχνολογίες (πλατφόρμες) αλλά επεξεργάστηκαν με τον καλύτερο δυνατό τρόπο για την εξάλειψη ανομοιογενειών. Πιθανοί λόγοι που οι μέθοδοι διόρθωσης πολλαπλών ερωτημάτων έπειτα από την επεξεργασία στο STATA δεν ανέδειξαν κανένα στατιστικώς σημαντικά διαφορικά εκφραζόμενο γονίδιο είναι ο μικρός αριθμός μελετών που χρησιμοποιήθηκαν εφόσον χρησιμοποιήσαμε το μοντέλο των τυχαίων επιδράσεων. Θα μπορούσε να επιλεγεί το μοντέλο των σταθερών επιδράσεων

κάνοντας μια πιο προσεκτική επιλογή των πειραματικών δεδομένων και φροντίζοντας για την πλήρη εξάλειψη των ανομοιογενειών σε αυτά.

Σημαντικό ρόλο διαδραματίζει η ανομοιογένεια των ιστών που χρησιμοποιήθηκαν στις σειρές δεδομένων καθώς οδήγησαν στο να συγκρίνουμε την γονιδιακή έκφραση σε κύτταρα αίματος, δέρματος και προστάτη. Επίσης το μεγαλύτερο μέρος των δειγμάτων που χρησιμοποιήθηκαν, προήλθαν από καλλιέργειες κυτταρικών σειρών στις οποίες έχει διαταραχθεί ο κυτταρικός κύκλος, ο οποίος και διαδραματίζει κυρίαρχο ρόλο στην απόκριση σε ιοντίζουσα ακτινοβολία.

Σε μελλοντικές έρευνες θα πρέπει να αξιοποιηθούν περισσότερα καινοτόμα και προσεκτικά σχεδιασμένα πειράματα για την αντιμετώπιση των σημαντικών ζητημάτων που σχετίζονται με την επαγωγή, την επεξεργασία και τη βιολογική σημασία της ακτινοβολίας. Επίσης οι μελέτες θα πρέπει να επικεντρωθούν στους βασικούς παράγοντες που καθορίζουν την απόκριση του οργανισμού: το οξειδωτικό στρες και τη φλεγμονή. Καταλήγοντας, σκοπός της κατανόησης των μηχανισμών που προκύπτουν από την διαφορική έκφραση των γονιδίων έπειτα από έκθεση σε ιοντίζουσα ακτινοβολία είναι αφενός η κατανόηση του τρόπου με τον οποίο γεννιέται κάποια ασθένεια ή κάποιος τύπος καρκίνου λόγω των χαμηλών επιπέδων ραδιενέργειας που δέχεται ένας οργανισμός καθημερινά ώστε να υπάρξει η κατάλληλη πρόληψη και αφετέρου η ικανοποιητική θεραπεία και βελτίωση της ποιότητας ζωής ασθενών που πρόκειται να θεραπευτούν με χρήση ιοντιζουσών ακτινοβολιών.

8 ΒΙΒΛΙΟΓΡΑΦΙΑ

- Aken, B.L., Achuthan, P., Akanni, W., Amode, M.R., Bernsdorff, F., Bhai, J., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., *et al.* (2017). Ensembl 2017. *Nucleic Acids Res* *45*, D635-D642.
- Alanazi, I.O., and Khan, Z. (2016). Understanding EGFR Signaling in Breast Cancer and Breast Cancer Stem Cells: Overexpression and Therapeutic Implications. *Asian Pac J Cancer Prev* *17*, 445-453.
- 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.
- Auvinen, A., Hakama, M., and Rytomaa, T. (1998). [Low-dose ionising radiation and cancer risk]. *Duodecim* *114*, 505, 507.
- Bagkos, P.G. (2015). Βιοπληροφορική (Greece: ΣΕΑΔ).
- Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., *et al.* (2011). NCBI GEO: archive for functional genomics data sets--10 years on. *Nucleic Acids Res* *39*, D1005-1010.
- 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.
- Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Sayers, E.W. (2013). GenBank. *Nucleic Acids Res* *41*, D36-42.
- 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.
- Borenstein, M., Hedges, L.V., Higgins, J.P.T., and Rothstein, H.R. (2009). *Introduction to Meta-Analysis* (UK).
- 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.
- Brown, P.O., and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nat Genet* *21*, 33-37.
- Butte, A. (2002). The use and analysis of microarray data. *Nat Rev Drug Discov* *1*, 951-960.
- 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.
- Chen, P.P.-S. (1976). The entity-relationship model--toward a unified view of data. *ACM Trans Database Syst* *1*, 9-36.
- Cheung, V.G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R., and Childs, G. (1999). Making and reading microarrays. *Nat Genet* *21*, 15-19.

- Choi, J.K., Yu, U., Kim, S., and Yoo, O.J. (2003). Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* *19 Suppl 1*, i84-90.
- Cleveland, W.S., Grosse, E., and Shyu, W.M. (1992). Local regression models. In *Statistical Models in S*, J.M. Chambers, and T.J. Hastie, eds. (London: Chapman & Hall).
- Clough, E., and Barrett, T. (2016). The Gene Expression Omnibus Database. *Methods Mol Biol* *1418*, 93-110.
- 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.
- Darwen, H. (2014). *An Introduction to Relational Database Theory*, 4th edn.
- Dauer, L.T., Brooks, A.L., Hoel, D.G., Morgan, W.F., Stram, D., and Tran, P. (2010). Review and evaluation of updated research on the health effects associated with low-dose ionising radiation. *Radiat Prot Dosimetry* *140*, 103-136.
- de Toledo, S.M., Buonanno, M., Li, M., Asaad, N., Qin, Y., Gonon, G., Shim, G., Galdass, M., Boateng, Y., Zhang, J., *et al.* (2011). The impact of adaptive and non-targeted effects in the biological responses to low dose/low fluence ionizing radiation: the modulating effect of linear energy transfer. *Health Phys* *100*, 290-292.
- Dhammi, I.K., and Kumar, S. (2014). Medical subject headings (MeSH) terms. *Indian J Orthop* *48*, 443-444.
- Du, P., Kibbe, W.A., and Lin, S.M. (2007). nuID: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. *Biol Direct* *2*, 16.
- Dudoit, S., and Speed, T.P. (2000). A score test for the linkage analysis of qualitative and quantitative traits based on identity by descent data from sib-pairs. *Biostatistics* *1*, 1-26.
- Dudoit, S., Yang, Y.H., Callow, M.J., and Speed, T.P. (2002). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica* *12*, 111-140.
- Duggan, D.J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J.M. (1999). Expression profiling using cDNA microarrays. *Nat Genet* *21*, 10-14.
- Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* *21*, 3439-3440.
- 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.
- Formann, A.K. (2010). The Newcomb-Benford law in its relation to some common distributions. *PLoS One* *5*, e10541.
- Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N., and Stratton, M.R. (2004). A census of human cancer genes. *Nat Rev Cancer* *4*, 177-183.
- Gene Ontology, C. (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res* *43*, D1049-1056.
- Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* *5*, R80.

- Georgakilas, A.G. (2008). Processing of DNA damage clusters in human cells: current status of knowledge. *Mol Biosyst* 4, 30-35.
- 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., Bennett, P.V., Wilson, D.M., 3rd, and Sutherland, B.M. (2004). Processing of bistranded abasic DNA clusters in gamma-irradiated human hematopoietic cells. *Nucleic Acids Res* 32, 5609-5620.
- Georgakilas, A.G., Martin, O.A., and Bonner, W.M. (2017). p21: A Two-Faced Genome Guardian. *Trends Mol Med* 23, 310-319.
- 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., Pavlopoulou, A., Louka, M., Nikitaki, Z., Vorgias, C.E., Bagos, P.G., and Michalopoulos, I. (2015). Emerging molecular networks common in ionizing radiation, immune and inflammatory responses by employing bioinformatics approaches. *Cancer Lett*.
- Ghandhi, S.A., Yaghoubian, B., and Amundson, S.A. (2008). Global gene expression analyses of bystander and alpha particle irradiated normal human lung fibroblasts: synchronous and differential responses. *BMC Med Genomics* 1, 63.
- Giglia, E. (2011). Quertle and KNALIJ: searching PubMed has never been so easy and effective. *Eur J Phys Rehabil Med* 47, 687-690.
- Goodhead, D.T., and Nikjoo, H. (1989). Track Structure Analysis of Ultrasoft X-rays Compared to High- and Low-LET Radiations. *International Journal of Radiation Biology* 55, 513-529.
- Gray, K.A., Yates, B., Seal, R.L., Wright, M.W., and Bruford, E.A. (2015). Genenames.org: the HGNC resources in 2015. *Nucleic Acids Res* 43, D1079-1085.
- Gruel, G., Voisin, P., Vaurijoux, A., Roch-Lefevre, S., Gregoire, E., Maltere, P., Petat, C., Gidrol, X., Voisin, P., and Roy, L. (2008). Broad modulation of gene expression in CD4+ lymphocyte subpopulations in response to low doses of ionizing radiation. *Radiat Res* 170, 335-344.
- Hada, M., and Georgakilas, A.G. (2008). Formation of clustered DNA damage after high-LET irradiation: a review. *J Radiat Res* 49, 203-210.
- Harris, S.L., and Levine, A.J. (2005). The p53 pathway: positive and negative feedback loops. *Oncogene* 24, 2899-2908.
- Hatzi, V.I., Laskaritou, D.A., Mavragani, I.V., Nikitaki, Z., Mangelis, A., Panayiotidis, M.I., Pantelias, G.E., Terzoudi, G.I., and Georgakilas, A.G. (2015). Non-targeted radiation effects in vivo: a critical glance of the future in radiobiology. *Cancer Lett* 356, 34-42.
- Held, K.D., Kawamura, H., Kaminuma, T., Paz, A.E., Yoshida, Y., Liu, Q., Willers, H., and Takahashi, A. (2016). Effects of Charged Particles on Human Tumor Cells. *Front Oncol* 6, 23.
- Heller, M.J. (2002). DNA microarray technology: devices, systems, and applications. *Annu Rev Biomed Eng* 4, 129-153.
- Hoaglin, D., Mosteller, F., and Tukey, J. (1983). *Understanding Robust and Exploratory Data Analysis* (New York: John Wiley and Sons, Inc.).
- Holland, B.S., and Copenhaver, M.D. (1987). An Improved Sequentially Rejective Bonferroni Test Procedure. *Biometrics* 43, 417-423.
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* 6, 65-70.

Hou, J., Wang, F., Kong, P., Yu, P.K., Wang, H., and Han, W. (2015). Gene profiling characteristics of radioadaptive response in AG01522 normal human fibroblasts. *PLoS One* *10*, e0123316.

Huber, P.J., and Ronchetti, E.M. (2009). *Robust Statistics*, Vol Hoboken, New Jersey

Canada, 2nd edn (John Wiley & Sons, Inc).

IAEA, I.A.E.A. (2004). *Radiation, People and the Environment* (Austria: Division of Radiation and Waste Safety).

International Atomic Energy Agency (IAEA) (2010). *Radiation Biology: A Handbook for Teachers and Students* (Vienna).

Jacoby, W.G. (2000). Loess: a nonparametric, graphical tool for depicting relationships between variables. *Electoral Studies* *19*, 577–613.

Kadhim, M., Salomaa, S., Wright, E., Hildebrandt, G., Belyakov, O.V., Prise, K.M., and Little, M.P. (2013). Non-targeted effects of ionising radiation--implications for low dose risk. *Mutat Res* *752*, 84-98.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* *28*, 27-30.

Kasprzyk, A. (2011). BioMart: driving a paradigm change in biological data management. *Database (Oxford)* *2011*, bar049.

Keith, S., Doyle, J.R., Harper, C., Mumtaz, M., Tarrago, O., Wohlers, D.W., Diamond, G.L., Citra, M., and Barber, L.E. (2012). Toxicological Profile for Radon. In *Toxicological Profile for Radon* (Atlanta (GA)).

Kelder, T., van Iersel, M.P., Hanspers, K., Kutmon, M., Conklin, B.R., Evelo, C.T., and Pico, A.R. (2012). WikiPathways: building research communities on biological pathways. *Nucleic Acids Res* *40*, D1301-1307.

Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. *Genome Res* *12*, 656-664.

Kitagawa, R., and Kastan, M.B. (2005). The ATM-dependent DNA damage signaling pathway. *Cold Spring Harb Symp Quant Biol* *70*, 99-109.

Kjellberg, R.N., Hanamura, T., Davis, K.R., Lyons, S.L., and Adams, R.D. (1983). Bragg-Peak Proton-Beam Therapy for Arteriovenous Malformations of the Brain. *New England Journal of Medicine* *309*, 269-274.

Koch, V., Majumder, A., and Wang, X.-N. (2006). Cherenkov Radiation from Jets in Heavy-Ion Collisions. *Physical Review Letters* *96*, 172302.

Lagarde, F. (2003). Methodology issues in epidemiological assessment of health effects of low-dose ionising radiation. *Radiat Prot Dosimetry* *104*, 297-314.

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.

Leroy, C., and Rancoita, P.-G. (2012). *Principles of radiation interaction in matter and detection* (New Jersey: World Scientific).

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* *1*, 417-425.

- Little, M.P. (2003). Risks associated with ionizing radiation. *Br Med Bull* 68, 259-275.
- Little, M.P., Wakeford, R., Tawn, E.J., Bouffler, S.D., and Berrington de Gonzalez, A. (2009). Risks associated with low doses and low dose rates of ionizing radiation: why linearity may be (almost) the best we can do. *Radiology* 251, 6-12.
- Lodermann, B., Wunderlich, R., Frey, S., Schorn, C., Stangl, S., Rodel, F., Keilholz, L., Fietkau, R., Gaipl, U.S., and Frey, B. (2012). Low dose ionising radiation leads to a NF-kappaB dependent decreased secretion of active IL-1beta by activated macrophages with a discontinuous dose-dependency. *Int J Radiat Biol* 88, 727-734.
- Long, X.H., Zhao, Z.Q., He, X.P., Wang, H.P., Xu, Q.Z., An, J., Bai, B., Sui, J.L., and Zhou, P.K. (2007). Dose-dependent expression changes of early response genes to ionizing radiation in human lymphoblastoid cells. *Int J Mol Med* 19, 607-615.
- 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.
- Low, F.E. (1958). Bremsstrahlung of Very Low-Energy Quanta in Elementary Particle Collisions. *Physical Review* 110, 974-977.
- Lyng, H., Badiie, A., Svendsrud, D.H., Hovig, E., Myklebost, O., and Stokke, T. (2004). Profound influence of microarray scanner characteristics on gene expression ratios: analysis and procedure for correction. *BMC Genomics* 5, 10.
- Martin, L.M., Marples, B., Lynch, T.H., Hollywood, D., and Maignol, L. (2014). Exposure to low dose ionising radiation: Molecular and clinical consequences. *Cancer Lett* 349, 98-106.
- McAulay, I.R., and Morgan, D. (1988). Natural Radioactivity in Soil in the Republic of Ireland.
- 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.
- Mezentsev, A., and Amundson, S.A. (2011). Global gene expression responses to low- or high-dose radiation in a human three-dimensional tissue model. *Radiat Res* 175, 677-688.
- Mi, H., Muruganujan, A., Casagrande, J.T., and Thomas, P.D. (2013). Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* 8, 1551-1566.
- Moreau, Y., Aerts, S., De Moor, B., De Strooper, B., and Dabrowski, M. (2003). Comparison and meta-analysis of microarray data: from the bench to the computer desk. *Trends Genet* 19, 570-577.
- National Research Council (2006). Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2 (Washington, D.C.: The National Academies Press).
- Normand, S.L. (1999). Meta-analysis: formulating, evaluating, combining, and reporting. *Stat Med* 18, 321-359.
- Oliver, S. (2003). On the MIAME Standards and Central Repositories of Microarray Data. *Comp Funct Genomics* 4, 1.
- Pagano, M. (1992). The histogram and boxplot: a picture is worth a thousand words. *Nutrition* 8, 374-375.
- Paul, S., and Amundson, S.A. (2008). Development of gene expression signatures for practical radiation biodosimetry. *Int J Radiat Oncol Biol Phys* 71, 1236-1244.

Pearson, W.R. (1994). Using the FASTA program to search protein and DNA sequence databases. *Methods Mol Biol* 24, 307-331.

PHP Documentation Group (1997). PHP.

Pico, A.R., Kelder, T., van Iersel, M.P., Hanspers, K., Conklin, B.R., and Evelo, C. (2008). WikiPathways: pathway editing for the people. *PLoS Biol* 6, e184.

Porter, L.E. (1985). Bethe-bloch stopping power parameters for light projectiles at energies near the stopping power maximum. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* 12, 50-55.

Qiu, X., Wu, H., and Hu, R. (2013). The impact of quantile and rank normalization procedures on the testing power of gene differential expression analysis. *BMC Bioinformatics* 14, 124.

Ramasamy, A., Mondry, A., Holmes, C.C., and Altman, D.G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med* 5, e184.

Ramos, F., and Wiame, J.M. (1982). Occurrence of a catabolic L-serine (L-threonine) deaminase in *Saccharomyces cerevisiae*. *Eur J Biochem* 123, 571-576.

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.

Riley, P.A. (1994). Free radicals in biology: oxidative stress and the effects of ionizing radiation. *Int J Radiat Biol* 65, 27-33.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43, e47.

Ritchie, M.E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A., and Smyth, G.K. (2007). A comparison of background correction methods for two-colour microarrays. *Bioinformatics* 23, 2700-2707.

Roberts, J.A., Hann, I.-H., and Slaughter, S.A. (2006). Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects. *Management Science* 52, 984-999.

Royeen, C.B. (1986). The boxplot: a screening test for research data. *Am J Occup Ther* 40, 569-571.

RStudio (2012). RStudio (Boston, MA).

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.

Samuels, M.L., Jeffrey A.Witmer, J.A., and Schaffner, A.A. (2012). *Statistics for the Life Sciences*, 4th edn (USA: Pearson Education, Inc.).

Schena, F.P. (2002). [Research and patents in the biomedical field]. *G Ital Nefrol* 19, 2-3.

Schena, M. (1996). Genome analysis with gene expression microarrays. *Bioessays* 18, 427-431.

Schuler, G., and Gutzwiller, F. (1991). [Low dose ionising radiation and cancer: findings and methods. Report of a meeting and consequences for Switzerland]. *Soz Praventivmed* 36, 209-216.

Šidák, Z. (1967). Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association*, 626-633.

- Silver, J.D., Ritchie, M.E., and Smyth, G.K. (2009). Microarray background correction: maximum likelihood estimation for the normal-exponential convolution. *Biostatistics* 10, 352-363.
- Sinden, R.R. (1994). *DNA Structure and Function* (USA: Academic Press, Inc.).
- 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.
- Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., and Kasprzyk, A. (2009). BioMart--biological queries made easy. *BMC Genomics* 10, 22.
- Smith, R., and Kao, G.D. (2004). Pictures, progress, and perplexities: the immediate cell biological effects of ionizing radiation. *Cancer Biol Ther* 3, 602-607.
- Smyth, G.K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3, 3.
- Smyth, G.K. (2005). Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (New York), pp. 397-420.
- Smyth, G.K., and Altman, N.S. (2013). Separate-channel analysis of two-channel microarrays: recovering inter-spot information. *BMC Bioinformatics* 14, 165.
- Smyth, G.K., Ritchie, M., Thorne, N., Wettenhall, J., Shi, W., Hu, Y., Bioinformatics Division, T.W.a.E.H.I., and of Medical Research, M., Australia (2002). *limma: Linear Models for Microarray and RNA-Seq Data_User's Guide*
(Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia
).
- Smyth, G.K., and Speed, T. (2003). Normalization of cDNA microarray data. *Methods* 31, 265-273.
- Stankevicius, L., Almeida da Silva, A.P., Ventura Dos Passos, F., Dos Santos Ferreira, E., Menks Ribeiro, M.C., M, G.D., E, J.P., Ferreira-Machado, S.C., Vassetzky, Y., de Almeida, C.E., *et al.* (2013). MiR-34a is up-regulated in response to low dose, low energy X-ray induced DNA damage in breast cells. *Radiat Oncol* 8, 231.
- StataCorp (2013). *Stata Statistical Software: Release 13* (College Station, TX: StataCorp LP).
- Stewart, R.D., Yu, V.K., Georgakilas, A.G., Koumenis, C., Park, J.H., and Carlson, D.J. (2011). Effects of radiation quality and oxygen on clustered DNA lesions and cell death. *Radiat Res* 176, 587-602.
- Sykes, P.J., and Day, T.K. (2007). Requirements for identification of low dose and non-linear mutagenic responses to ionising radiation. *Dose Response* 5, 308-314.
- Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., *et al.* (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 43, D447-452.
- Thakkinstian, A., McElduff, P., D'Este, C., Duffy, D., and Attia, J. (2005). A method for meta-analysis of molecular association studies. *Stat Med* 24, 1291-1306.
- Toprani, S.M., and Das, B. (2015). Role of base excision repair genes and proteins in gamma-irradiated resting human peripheral blood mononuclear cells. *Mutagenesis* 30, 247-261.

- Townsend, J.P. (2004). Resolution of large and small differences in gene expression using models for the Bayesian analysis of gene expression levels and spotted DNA microarrays. *BMC Bioinformatics* 5, 54.
- UNEP, U.N.E.P. (2016). Radiation Effects and Sources (Austria: United Nations Environment Programme).
- Valentin, J. (2005). Low-dose extrapolation of radiation-related cancer risk. *Ann ICRP* 35, 1-140.
- Wallace, S.S. (1998). Enzymatic processing of radiation-induced free radical damage in DNA. *Radiat Res* 150, S60-79.
- Wang, H.P., Long, X.H., Sun, Z.Z., Rigaud, O., Xu, Q.Z., Huang, Y.C., Sui, J.L., Bai, B., and Zhou, P.K. (2006). Identification of differentially transcribed genes in human lymphoblastoid cells irradiated with 0.5 Gy of gamma-ray and the involvement of low dose radiation inducible CHD6 gene in cell proliferation and radiosensitivity. *Int J Radiat Biol* 82, 181-190.
- Wang, J., Duncan, D., Shi, Z., and Zhang, B. (2013). WEB-based GENE SeT Analysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res* 41, W77-83.
- Wang, X. (2008). miRDB: a microRNA target prediction and functional annotation database with a wiki interface. *RNA* 14, 1012-1017.
- Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., *et al.* (2016). Ensembl 2016. *Nucleic Acids Res* 44, D710-716.
- Yu, X., and Yi, C. (2010). Design and Implementation of the Website Based on PHP & MySQL. Paper presented at: 2010 International Conference on E-Product E-Service and E-Entertainment.
- Zar, J.H. (2010). *Biostatistical Analysis*, 5th edn (USA: Pearson Education, Inc.).
- 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.
- Zhou, D.S. (1980). [An approach to the statistical treatment of multiple groups of small frequencies in medical research--hypergeometric probability distribution and its application (author's transl)]. *Zhonghua Yu Fang Yi Xue Za Zhi* 14, 211-213.
- Ziegler, J.F., Ziegler, M.D., and P., B.J. (2010). SRIM – The stopping and range of ions in matter. Science Direct.