



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών
Τομέας Φυσικής - Εργαστήριο βλαβών DNA

Διπλωματική Εργασία

*«Μελέτη της απόκρισης ανθρώπινων
κυττάρων στις σύνθετες βλάβες DNA
με χρήση μεθόδων βιοπληροφορικής»*

Βαγγέλης Γιουκάκης

Επιβλέπων:
Δρ. Αλέξανδρος Γεωργακίλας,
Αν. καθηγητής ΣΕΜΦΕ ΕΜΠ

Αθήνα, 2020



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών
Τομέας Φυσικής – Εργαστήριο βλαβών DNA

**«Μελέτη της απόκρισης ανθρώπινων
κυττάρων στις σύνθετες βλάβες DNA
με χρήση μεθόδων βιοπληροφορικής»**

Διπλωματική εργασία
του

Βαγγέλη Γιουκάκη

Επιβλέπων: Δρ. Αλέξανδρος Γεωργακίλας, Αν. Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την Τρίτη 13 Οκτωβρίου 2020

.....

Αλέξανδρος Γεωργακίλας

Αν. Καθηγητής ΕΜΠ

.....

Απόστολος Κυρίτης

Αν. Καθηγητής ΕΜΠ

.....

Κων/νος Αναγνωστόπουλος

Αν. Καθηγητής ΕΜΠ

Περίληψη

Σκοπός της διπλωματικής εργασίας είναι η μελέτη της απόκρισης που παρουσιάζουν τα ανθρώπινα κύτταρα, όταν προκληθούν σύνθετες βλάβες στο γενετικό υλικό που βρίσκεται στο εσωτερικό τους. Οι σύνθετες ή ομαδοποιημένες βλάβες του γενετικού υλικού αποτελούν μεγάλο πρόβλημα για τα ανθρώπινα (και όχι μόνο) κύτταρα και συχνά οδηγούν σε μεταλλάξεις και κυτταρικό θάνατο. Η βιολογική τους σημασία έγκειται στην δυσκολία των κυττάρων να τις επιδιορθώσουν με επιτυχία σε αντίθεση με τις μεμονωμένες βλάβες που επιδιορθώνονται σχετικά εύκολα από τους επιδιορθωτικούς μηχανισμούς του DNA. Στο πλαίσιο της παρούσας διπλωματικής εργασίας χρησιμοποιήθηκαν μέθοδοι βιοπληροφορικής για τον εντοπισμό όλων των γονιδίων που συμμετέχουν στην αντίδραση των ανθρώπινων κυττάρων όταν στο γενετικό τους υλικό προκληθούν σύνθετες βλάβες. Στην συνέχεια έγινε φιλτράρισμα των αποτελεσμάτων ώστε να προκύψουν τα γονίδια εκείνα που έχουν την μεγαλύτερη πιθανότητα να συμμετέχουν όντως στην απόκριση των κυττάρων στις σύνθετες βλάβες DNA. Τα γονίδια που προέκυψαν μετά το φιλτράρισμα αναλύθηκαν περαιτέρω με την χρήση τεσσάρων διαφορετικών εργαλείων βιοπληροφορικής ανάλυσης για να επιβεβαιωθεί η ενεργοποίηση τους μετά την πρόκληση σύνθετων βλαβών στο DNA. Συγκεκριμένα ελέγχθηκαν οι βιολογικοί μηχανισμοί στους οποίους συμμετέχουν, η συσχέτιση τους με διάφορες ασθένειες, η επίδραση που έχουν στην επιβίωση των ασθενών σε διάφορους τύπους καρκίνου και η συσχέτιση τους με την κατάσταση της κυτταρικής γήρανσης. Από τους βιολογικούς μηχανισμούς που συμμετέχουν τα γονίδια που αναλύθηκαν, βρέθηκαν εννέα που υποδεικνύουν την ύπαρξη σύνθετων βλαβών DNA. Για τα περισσότερα γονίδια υπάρχει συσχέτιση τους με ασθένειες και τύπους καρκίνου ενώ επίσης για τα περισσότερα από αυτά υπάρχουν αναφορές για συσχέτιση με την κυτταρική γήρανση. Τέλος επισημάνθηκαν 15 γονίδια ως τα πλέον πιθανά να ενεργοποιούνται με την ύπαρξη σύνθετων βλαβών στο γενετικό υλικό των κυττάρων και τα οποία χρήζουν περαιτέρω έρευνας.

Abstract

The aim of this thesis is to study the response of human cells to the induction of complex DNA lesions. Complex or clustered DNA lesions are extremely dangerous for human cells (and those of other organisms) as they can lead to mutations and cell death. The biological significance of such lesions relates to the inability of cells to process them efficiently compared to isolated DNA lesions that are repaired relatively easy with the 'DNA repairing mechanisms'. In this thesis, bioinformatic methods were used to detect all genes that take part in the reaction of human cells when complex DNA damage is induced to them. A filtering process was conducted to the results in order to gather those genes that are the most probable to take part in the response of cells to complex DNA damage. The genes that were gathered after the filtering process were analyzed using four different bioinformatic platforms to confirm their regulation after the induction of clustered DNA damage. More specifically we checked the biological pathways that these genes belong to, their association with certain diseases, the effect they have on patient survival in different types of cancer and their association with cellular senescence. From the biological pathways that the genes that were analyzed belong to, we found nine that suggest the existence of complex DNA damage. For most of these genes there is correlation with diseases and types of cancer while again for most of them there are reports for connection with cellular senescence. Finally fifteen genes were pointed out as the most probable for being activated with complex DNA damage and as subjects for more thorough research.

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή μου κύριο Αλέξανδρο Γεωργακίλα που μου έδωσε την ευκαιρία να αποτελέσω μέρος της ερευνητικής του ομάδας στο εργαστήριο βλαβών DNA και να ασχοληθώ με ένα τόσο ενδιαφέρον θέμα καθώς και για την πολύτιμη καθοδήγηση του κατά την εκπόνηση της παρούσας διπλωματικής εργασίας.

Θα ήθελα επίσης να ευχαριστήσω θερμά την Δρ. Ζαχαρένια Νικητάκη δίχως την βοήθεια της οποίας θα ήταν αδύνατη η ολοκλήρωση της συγκεκριμένης διπλωματικής. Οι συμβουλές και οι παρατηρήσεις της καθώς και η άμεση ανταπόκριση της κάθε φορά που χρειάστηκα οποιαδήποτε βοήθεια, ήταν καθοριστικής σημασίας.

Ευχαριστώ πολύ την Δρ. Αθανασία Παυλοπούλου για την συμβολή της στο υπολογιστικό κομμάτι της εργασίας και την εκμάθηση της πλατφόρμας GEPIA που ήταν ένα από τα τέσσερα εργαλεία βιοπληροφορικής ανάλυσης που χρησιμοποιήθηκαν.

Ευχαριστώ επίσης την τριμελή επιτροπή που δέχτηκαν να συμμετάσχουν στην αξιολόγηση της διπλωματικής εργασίας.

Θέλω ακόμα να ευχαριστήσω τον αδερφό μου και τους φίλους μου που με την συντροφιά τους έκαναν την περίοδο εκπόνησης της διπλωματικής πιο εύκολη.

Τέλος θέλω να ευχαριστήσω τους γονείς μου που χωρίς την ανυπολόγιστη στήριξη τους σε όλους τους τομείς τίποτα από όλα αυτά δεν θα ήταν δυνατό.

Περιεχόμενα

Περίληψη.....	3
Abstract.....	4
Ευχαριστίες.....	5
Περιεχόμενα.....	7

Α' Θεωρητικό Μέρος

Κεφάλαιο 1: Στοιχεία βιολογίας και ακτινοβιολογίας.....12

1.1 Εισαγωγή.....	12
1.2 Το ευκαρυωτικό κύτταρο και το γενετικό υλικό	13
1.2.1 Δομή του ευκαρυωτικού κυττάρου	13
1.2.2 Το γενετικό υλικό (δομή του DNA).....	16
1.3 Ο κυτταρικός κύκλος και η Απόπτωση.....	18
1.4 Ελαγωγή βλαβών DNA.....	20
1.4.1 Οξειδωτικό στρες.....	20
1.4.2 Ακτινοβολίες.....	20
1.4.3 Στρες αντιγραφής.....	21
1.5 Ιοντίζουσες ακτινοβολίες.....	22
1.5.1 Εισαγωγή.....	22
1.5.2 Τυπικές ιοντίζουσες ακτινοβολίες.....	23
1.5.3 Στάδια αλληλεπίδρασης ακτινοβολίας και βιολογικής ύλης.....	24
1.5.4 Άμεσα και έμμεσα αποτελέσματα δράσης της ακτινοβολίας.....	25

Κεφάλαιο 2: Βλάβες DNA.....27

2.1 Η Γραμμική Εναπόθεση Ενέργειας και πυκνότητα βλαβών.....	27
2.2 Είδη βλαβών DNA.....	28
2.2.1 Μεμονωμένες βλάβες.....	28
2.2.2 Ομαδοποιημένες βλάβες.....	29
2.3 Επιδιορθωτικοί μηχανισμοί.....	30
2.3.1 Εισαγωγή.....	30
2.3.2 Μηχανισμός Εκτομής Βάσης (BER).....	31
2.3.3 Μηχανισμός Εκτομής Νουκλεοτιδίων (NER).....	32

2.3.4	Επιδιόρθωση αναντιστοιχίας (MMR).....	35
2.3.5	Ομόλογος ανασυνδυασμός (HR).....	37
2.3.6	Μη-Ομόλογη Ένωση Άκρων (NHEJ).....	40
Κεφάλαιο 3: Σύνθετες βλάβες DNA.....		43
3.1	Η βιολογική σημασία των σύνθετων βλαβών.....	43
3.2	Μέθοδοι ανίχνευσης σύνθετων βλαβών DNA.....	44
3.3	Σύνθετες βλάβες DNA και συστημικά φαινόμενα.....	46
Κεφάλαιο 4 Στοιχεία Βιοπληροφορικής.....		48
4.1	Εισαγωγή - Η σημασία της βιοπληροφορικής.....	48
4.2	Όροι γονιδιακής οντολογίας.....	49
4.3	Βάσεις δεδομένων.....	51
4.4	Πλατφόρμες Βιοπληροφορικής Ανάλυσης.....	52
4.4.1	Database for Annotation, Visualization and Integrated Discovery (DAVID).....	52
4.4.2	STRING.....	53
4.4.3	DisGeNET.....	54
4.4.4	Bioinformatics and Evolutionary Genomics (Van de Peer Lab.....	55
4.4.5	Gene Expression Profiling Interactive Analysis (GEPIA).....	55
4.4.6	SeneQuest.....	56
Β' Υπολογιστικό Μέρος		
Κεφάλαιο 5: Εύρεση γονιδίων που συμμετέχουν στις		
σύνθετες βλάβες του DNA.....		60
5.1	Σκοπός.....	60
5.2	Μεθοδολογία.....	61
5.3	Αποτελέσματα.....	64
5.4	Συνοψιση αποτελεσμάτων - Συμπεράσματα.....	73
Κεφάλαιο 6: Περαιτέρω διερεύνηση των γονιδίων που		
επισημάνθηκαν ως κρίσιμα.....		75
6.1	Βιολογικοί μηχανισμοί.....	75
6.1.1	Στόχος.....	75
6.1.2	Μέθοδος.....	76

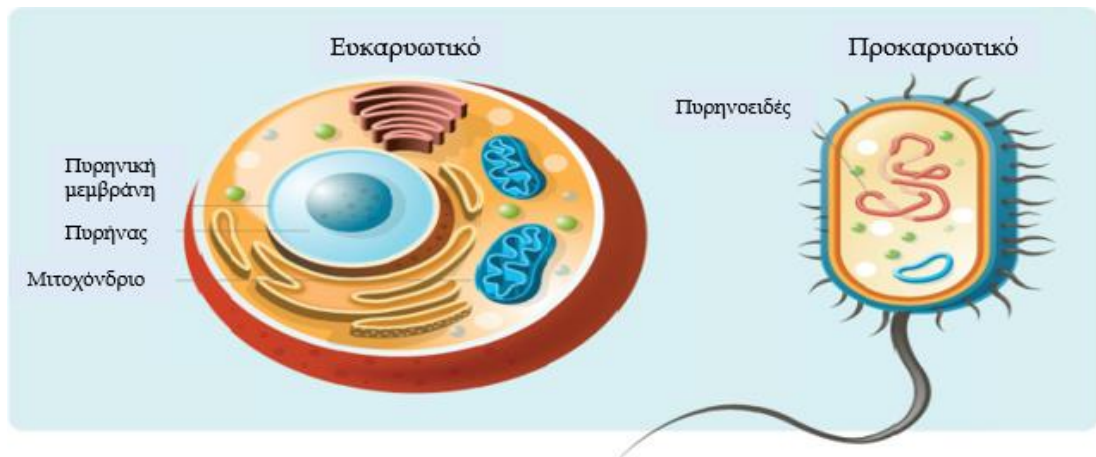
6.1.3 Αποτελέσματα.....	76
6.2 Ασθένειες και τύποι καρκίνου.....	80
6.2.1 Στόχος.....	80
6.2.2 Μέθοδος.....	80
6.2.3 Αποτελέσματα.....	81
6.3 Κυτταρική γήρανση (<i>senescence</i>).....	87
6.3.1 Στόχος.....	87
6.3.2 Μέθοδος.....	87
6.3.3 Αποτελέσματα.....	88
6.4 Συνόψιση αποτελεσμάτων – συμπεράσματα.....	89
Κεφάλαιο 7: Συμπεράσματα της διπλωματικής εργασίας και μελλοντικές κατευθύνσεις.....	91
Βιβλιογραφία.....	94

Α' Θεωρητικό Μέρος

Κεφάλαιο 1: Στοιχεία βιολογίας και ακτινοβιολογίας

1.1 Εισαγωγή

Σε αυτό το κεφάλαιο θα παρουσιάσουμε μερικά από τα βασικά στοιχεία της κυτταρικής βιολογίας. Η κυτταρική βιολογία η οποία παλιότερα ήταν γνωστή με τον όρο κυτταρολογία, είναι κλάδος της βιολογίας που μελετάει την δομή και την λειτουργία του κυττάρου, ενώ η έρευνα και οι ανακαλύψεις στον συγκεκριμένο κλάδο έχουν συμβάλλει σε πολύ μεγάλο βαθμό στην πρόοδο της ιατρικής και στην εύρεση μεθόδων για την καταπολέμηση του καρκίνου. Το κύτταρο αποτελεί την θεμελιώδη δομική και λειτουργική μονάδα των ζωντανών οργανισμών ενώ έχει υπολογιστεί ότι το μέσο ανθρώπινο σώμα αποτελείται από περίπου 40 τρισεκατομμύρια κύτταρα [1]. Οι οργανισμοί μπορούν να κατηγοριοποιηθούν σε μονοκύτταρους, δηλαδή αυτούς που αποτελούνται από ένα μόλις κύτταρο (όπως για παράδειγμα τα βακτήρια) και σε πολυκύτταρους, κατηγορία που περιλαμβάνει όλα τα φυτά και ζώα του πλανήτη. Με τις βιοχημικές διαδικασίες που πραγματοποιούνται στο εσωτερικό των κυττάρων παράγεται η απαραίτητη ενέργεια για την επιβίωση και εξέλιξη των έμβιων όντων. Τα κύτταρα διαχωρίζονται (βλέπε και εικόνα 1.1) σε προκαρυωτικά, δηλαδή αυτά που δεν έχουν πυρήνα στο εσωτερικό τους (συναντώνται στους προκαρυωτικούς - μονοκύτταρους οργανισμούς) και σε ευκαρυωτικά, δηλαδή αυτά που στο εσωτερικό τους βρίσκεται μια δομή που ονομάζεται πυρήνας [2]. Στην συγκεκριμένη εργασία και εφόσον το αντικείμενο της μελέτης μας είναι τα ανθρώπινα κύτταρα (πολυκύτταρος οργανισμός) θα ασχοληθούμε αποκλειστικά με τα ευκαρυωτικά κύτταρα.



Εικόνα 1.1: Το ευκαρυωτικό και το προκαρυωτικό κύτταρο

1.2 Το ευκαρυωτικό κύτταρο και το γενετικό υλικό

1.2.1 Δομή του ευκαρυωτικού κυττάρου

Ακόμα και τα ευκαρυωτικά κύτταρα δεν είναι όλα όμοια μεταξύ τους καθώς η μορφή και η δομή τους ποικίλλει ανάλογα με τον οργανισμό στον οποίο ανήκουν (πχ ζωικά και φυτικά ευκαρυωτικά κύτταρα) και την λειτουργία που επιτελεί το καθένα τους (πχ μυϊκά και νευρικά κύτταρα). Παρ'όλα αυτά κάποιες βασικές δομές απαντώνται σε όλους τους τύπους ευκαρυωτικών κυττάρων (βλέπε και εικόνα 1.2):

Η **κυτταρική μεμβράνη** διαχωρίζει και προστατεύει το κύτταρο από το εξωτερικό του περιβάλλον και αποτελείται κυρίως από ένα διπλό στρώμα πρωτεϊνών και λιπιδίων. Πρόκειται στην ουσία για το εξωτερικό σύνορο του κυττάρου, στο εσωτερικό του οποίου βρίσκεται το σύνολο των οργανιδίων που το αποτελούν. Ενσωματωμένα στην κυτταρική μεμβράνη βρίσκονται διάφορα μόρια τα οποία λειτουργούν σαν αγωγοί και αντλίες, επιτρέποντας την μετακίνηση μορίων μέσα και έξω από το κύτταρο [3].

Ο **κυτταρικός σκελετός** είναι μια σύνθετη και δυναμική δομή που αποτελείται κυρίως από πρωτεΐνες και χρησιμεύει στο να οργανώνει και να διατηρεί το σχήμα του κυττάρου, κρατώντας τα διάφορα οργανίδια στην θέση τους.

Το **κυτταρόπλασμα** είναι το υγρό που γεμίζει τον εσωτερικό χώρο του κυττάρου και αποτελείται κυρίως από νερό, μικρά μόρια, ιόντα και πρωτεΐνες. Χρησιμεύει στην υδρόλυση θρεπτικών ουσιών και αποβλήτων του κυτταρικού μεταβολισμού καθώς και στην μεταφορά ουσιών στο εσωτερικό του κυττάρου από το ένα οργανίδιο στο άλλο. Περιέχει επίσης άλατα και είναι εξαιρετικός αγωγός του ηλεκτρικού ρεύματος, δημιουργώντας ιδανικές συνθήκες για τις λειτουργίες του κυττάρου.

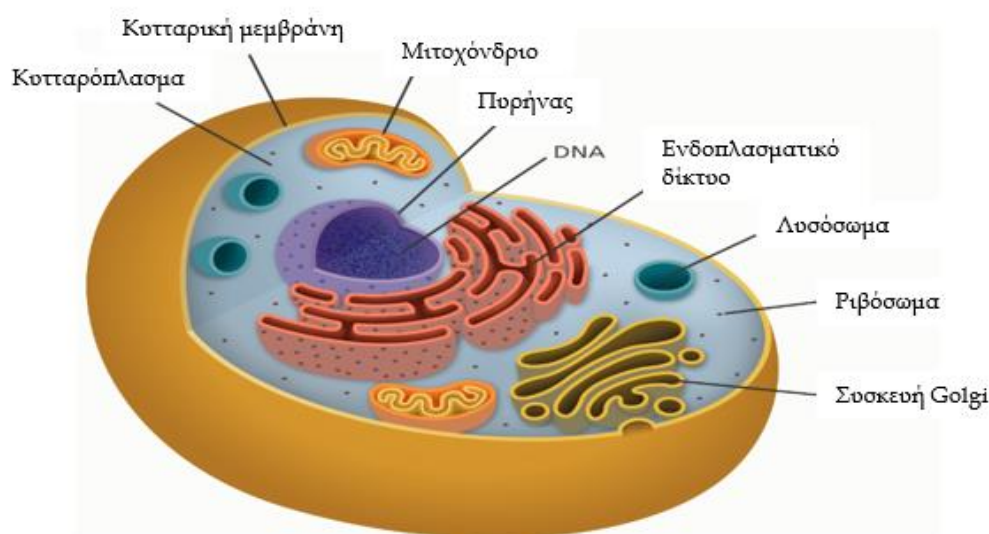
Ο **πυρήνας** είναι το πιο «διάσημο» οργανίδιο των ευκαρυωτικών κυττάρων. Περιέχει τα χρωμοσώματα του κυττάρου και είναι το μέρος που πραγματοποιείται η αντιγραφή του DNA και η σύνθεση του RNA. Ο πυρήνας έχει σφαιροειδές σχήμα και χωρίζεται από το κυτταρόπλασμα με μια μεμβράνη που ονομάζεται πυρηνικός φάκελος. Ο πυρηνικός φάκελος απομονώνει και προστατεύει το γενετικό υλικό (χρωμοσώματα) από διάφορα μόρια που θα μπορούσαν να βλάψουν την δομή του ή να παρεμποδίσουν την επεξεργασία του. Μέσα στον πυρήνα το DNA «μεταγράφεται» στο ειδικό μόριο RNA που ονομάζεται mRNA και στην συνέχεια μεταφέρεται έξω από τον πυρήνα, όπου «μεταφράζεται» σε μόριο πρωτεΐνης.

Τα **ριβοσώματα** είναι μεγάλα σύμπλοκα μορίων, περιλαμβανομένων πρωτεϊνών και μορίων RNA, τα οποία είναι υπεύθυνα για την μετατροπή της γενετικής πληροφορίας που φέρουν τα μόρια mRNA στην κατάλληλη αλληλουχία αμινοξέων που συνθέτουν ένα μόριο πρωτεΐνης. Η επεξεργασία αυτή της γενετικής πληροφορίας για την σύνθεση πρωτεϊνών ονομάζεται «μετάφραση». Η πρωτεΐνοσύνθεση είναι πολύ σημαντική διαδικασία για το κύτταρο και για τον λόγο αυτό στο εσωτερικό του υπάρχει μεγάλος αριθμός ριβοσωμάτων (εκατοντάδες ή ακόμα και χιλιάδες σε ορισμένες περιπτώσεις). Τα ριβοσώματα μετακινούνται ελεύθερα στο κυτταρόπλασμα ή κάποιες φορές προσδένονται σε ένα άλλο οργανίδιο που ονομάζεται **ενδοπλασματικό δίκτυο**. Μια πρωτεΐνη που προορίζεται να βγει έξω από το κύτταρο μετά την δημιουργία της από τα

ριβοσώματα, θα περάσει από ένα άλλο οργανίδιο που ονομάζεται **συσκευή Golgi** για επιπλέον επεξεργασία και πακετάρισμα.

Τα **μιτοχόνδρια** είναι οργανίδια που απαντώνται σε διάφορα μεγέθη, σχήματα και αριθμούς στο κυτταρόπλασμα όλων των ευκαρυωτικών κυττάρων. Τα μιτοχόνδρια περιέχουν το δικό τους γενετικό υλικό, το οποίο είναι ξεχωριστό από αυτό του πυρήνα καθώς και δύο ξεχωριστές μεμβράνες, μια στο εξωτερικό τους που περιβάλλει ολόκληρο το οργανίδιο και μια στο εσωτερικό που αναδιπλώνεται και προεξέχει. Τα μιτοχόνδρια παίζουν πολύ σημαντικό ρόλο στην παραγωγή ενέργειας για το κύτταρο μέσω του μεταβολισμού βιολογικών μορίων που προσλαμβάνονται από την τροφή. Για τον μεταβολισμό αυτό ακολουθούνται διάφορα βιολογικά «μονοπάτια» (pathways) όπως η «γλυκόλυση» (glycolysis).

Τα **λυσοσώματα** συχνά αναφέρονται ως το πεπτικό σύστημα του κυττάρου. Είναι σφαιρικά οργανίδια που περιφράσσονται από μια μεμβράνη και περιέχουν πολλά πεπτικά ένζυμα, δηλαδή πρωτεΐνες που επιταχύνουν τις βιοχημικές διαδικασίες. Τα ένζυμα αυτά χρησιμεύουν στο να διασπούν πρωτεΐνες, νουκλεϊκά οξέα και πολυσακχαρίτες. Επίσης τα λυσοσώματα «χωνεύουν», δηλαδή καταστρέφουν ξένα βακτήρια τα οποία εισέρχονται στο κύτταρο βοηθώντας στην προστασία του και αποσυνθέτουν μη λειτουργικά οργανίδια, όπως μιτοχόνδρια που έχουν σταματήσει να λειτουργούν [3].



Εικόνα 1.2: Το ευκαρυωτικό κύτταρο

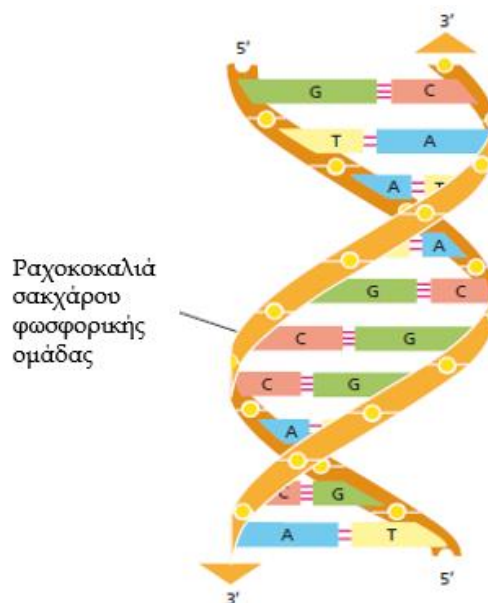
1.2.2 Το γενετικό υλικό (δομή του DNA)

Η ζωή εξαρτάται από την ικανότητα των κυττάρων να αποθηκεύουν και να ανακτούν τις γενετικές πληροφορίες που απαιτούνται για την κατασκευή και την συντήρηση ενός ζωντανού οργανισμού. Αυτή η πληροφορία περνάει από την μια γενιά κυττάρων στην επόμενη μέσω της κυτταρικής διαίρεσης και από την μια γενιά πολυκύτταρων οργανισμών στην επόμενη μέσω των αναπαραγωγικών κυττάρων, δηλαδή τα ωάρια και τα σπερματοζωάρια [4]. Οι γενετικές πληροφορίες σε όλες τις κυτταρικές μορφές ζωής αποθηκεύονται στα ειδικά μόρια που ονομάζονται δεοξυριβονουκλεϊκά οξέα ή αλλιώς DNA (Deoxyribonucleic Acid). Το σύνολο των μορίων DNA που υπάρχουν σε ένα κύτταρο αποτελούν το γενετικό υλικό του κυττάρου. Τα δεοξυριβονουκλεϊκά οξέα είναι βιολογικά πολυμερή ή αλλιώς μακρομόρια, δηλαδή σχετικά μεγάλα μόρια (μόρια με υψηλό μοριακό βάρος) που αποτελούνται από επαναλαμβανόμενες απλές μονάδες που ονομάζονται μονομερή. Στην περίπτωση του DNA τα μονομερή είναι τα νουκλεοτίδια (βλέπε εικόνα 1.3), σύνθετα οργανικά μόρια που αποτελούνται από τρία επιμέρους μόρια που συνδέονται μεταξύ τους με ομοιοπολικό δεσμό: μια πεντόζη-δεοξυριβόζη (σάκχαρο με πέντε άτομα άνθρακα), ένα μόριο φωσφορικού οξέος και μιας εκ των τεσσάρων αζωτούχων βάσεων: αδερίνη, γουανίνη, κυτοσίνη και θυμίνη.



Εικόνα 1.3: Ένα νουκλεοτίδιο αποτελείται από ένα σάκχαρο (δεοξυριβόζη), ένα μόριο φωσφορικού οξέος και μια αζωτούχα βάση (στην προκειμένη περίπτωση μια γουανίνη). Πολλά τέτοια νουκλεοτίδια συνδέονται για την δημιουργία μιας αλυσίδας.

Μια πολυνουκλεοτιδική αλυσίδα (εικόνα 1.3) σχηματίζεται από ένα σύνολο νουκλεοτιδίων τα οποία ενώνονται μεταξύ τους με ομοιοπολικό δεσμό που αναπτύσσεται μεταξύ του 3' άνθρακα του σακχάρου του ενός νουκλεοτιδίου με την φωσφορική ομάδα που βρίσκεται στον 5' άνθρακα του επόμενου νουκλεοτιδίου. Έτσι δημιουργείται μια ραχοκοκαλιά από επαναλαμβανόμενα μόρια σακχάρου-φωσφορικής ομάδας από τα οποία εξέχουν οι αζωτούχες βάσεις. Ένα μόριο DNA (εικόνα 1.4) σχηματίζεται από δύο πολυνουκλεοτιδικές αλυσίδες οι οποίες συνδέονται με δεσμούς υδρογόνου που σχηματίζονται ανάμεσα στις αζωτούχες βάσεις της κάθε μιας μέσω του κανόνα της συμπληρωματικότητας: η αδενίνη συνδέεται με την θυμίνη με διπλό δεσμό υδρογόνου και η γουανίνη με την κυτοσίνη με τριπλό δεσμό υδρογόνου [4]. Αυτή ακριβώς η αλληλουχία των αζωτούχων βάσεων σε ένα μόριο DNA (οργανωμένες σε λειτουργικές μονάδες που ονομάζονται γονίδια) περιέχει όλη την γενετική πληροφορία την οποία το κύτταρο χρησιμοποιεί με ειδικούς μηχανισμούς (*μεταγραφή και μετάφραση*) για την παραγωγή πρωτεϊνών που εκτελούν όλες τις λειτουργίες του οργανισμού. Τα μόρια του DNA συμπυκνώνονται («πακετάρονται») σε μεγάλο βαθμό με την βοήθεια επιπλέον πρωτεϊνών που ονομάζονται ιστόνες και καταλήγουν στην δομή των χρωμοσωμάτων που παρατηρούμε με το μικροσκόπιο.

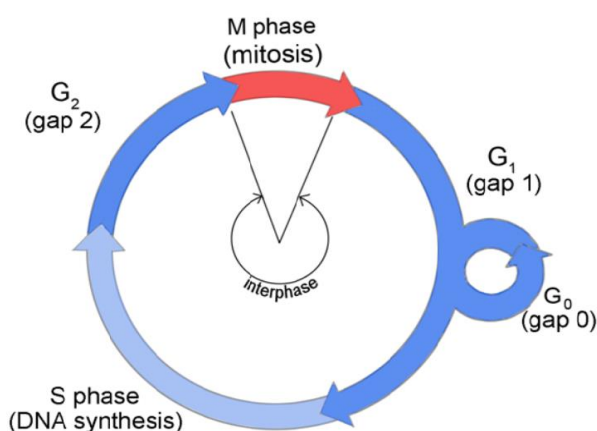


Εικόνα 1.4: Το μόριο του DNA

1.3 Ο κυτταρικός κύκλος και η Απόπτωση

Με τον όρο κυτταρικό κύκλο εννοούμε ουσιαστικά τον κύκλο ζωής του κυττάρου. Είναι μια σειρά από γεγονότα και λειτουργίες που λαμβάνουν χώρα στο κύτταρο και που έχουν σαν κατάληξη είτε την κυτταρική διαίρεση είτε τον θάνατο του κυττάρου. Οι λεπτομέρειες του κυτταρικού κύκλου μπορεί να διαφέρουν από οργανισμό σε οργανισμό, ενώ και η χρονική διάρκεια του κύκλου διαφέρει για διαφορετικά είδη κυττάρων στον ίδιο οργανισμό. Τα βασικά χαρακτηριστικά του κυτταρικού κύκλου όμως είναι καθολικά καθώς επιτρέπουν στο κύτταρο να επιτελεί την θεμελιώδη λειτουργία του που είναι η αντιγραφή και το πέρασμα της γενετικής πληροφορίας στην επόμενη γενιά κυττάρων [4]. Ο κυτταρικός κύκλος (εικόνα 1.5) μπορεί να χωριστεί σε δύο επιμέρους φάσεις: την μεσόφαση (interphase) και την μίτωση ενώ η μεσόφαση χωρίζεται περαιτέρω στις φάσεις G₁, S και G₂. Σε όλη την διάρκεια της μεσόφασης το κύτταρο αναπτύσσεται και επιτελεί τις λειτουργίες του, συνθέτοντας πρωτεΐνες και αυξάνοντας την μάζα του. Κατά την S φάση πραγματοποιείται η αντιγραφή του DNA που προορίζεται να δοθεί στα δύο θυγατρικά κύτταρα που θα προκύψουν από την διαίρεση του αρχικού. Από την S φάση προηγείται η G₁ ενώ μετά την S ακολουθεί η G₂. Σε αυτές τις φάσεις το κύτταρο, εκτός από τις βασικές λειτουργίες του ελέγχει τόσο την εσωτερική κατάσταση όσο και το εξωτερικό του περιβάλλον, στα λεγόμενα «σημεία ελέγχου», για να βεβαιωθεί ότι οι συνθήκες είναι κατάλληλες πριν προχωρήσει στα πολύ σημαντικά γεγονότα της αντιγραφής του DNA (μετά την G₁) και της μίτωσης (μετά την G₂). Με αυτό τον τρόπο εξασφαλίζεται ότι τα γεγονότα του κυτταρικού κύκλου θα γίνουν με μια καθορισμένη σειρά και ότι κάθε διαδικασία έχει ολοκληρωθεί πριν ξεκινήσει η επόμενη. Ανάλογα τώρα με το πρόβλημα που εντοπίζεται το κύτταρο έχει συγκεκριμένες μεθόδους να το αντιμετωπίζει. Αν για παράδειγμα εντοπιστεί βλάβη στο DNA ο κυτταρικός κύκλος διακόπτεται μέχρι να επιδιορθωθεί η βλάβη με κάποιον από τους γνωστούς μηχανισμούς επιδιόρθωσης (βλέπε παράγραφο 2.3). Στην ειδική περίπτωση που το κύτταρο αντιληφθεί ότι οι εξωτερικές συνθήκες δεν είναι κατάλληλες για την ανάπτυξη του, έχει την δυνατότητα είτε να επεκτείνει την διάρκεια της G₁ φάσης ή και να εισέλθει σε μια

ειδική κατάσταση αδράνειας που ονομάζεται G_0 . Ορισμένα κύτταρα μάλιστα βρίσκονται μόνιμα στην φάση G_0 μέχρι να τους δοθεί εντολή να ξεκινήσουν τον πολλαπλασιασμό τους, όπως για παράδειγμα τα κύτταρα του συκωτιού τα οποία είναι εν γένει αδρανή μέχρι να εντοπιστεί βλάβη στο συκώτι όπου ξεκινάνε να πολλαπλασιάζονται για να επιδιορθωθεί η ζημιά [4]. Μετά την G_2 φάση και αν το κύτταρο έχει περάσει όλους τους ελέγχους, ακολουθεί η M φάση δηλαδή η μίτωση κατά την οποία πραγματοποιείται η σταδιακή διαίρεση του κυττάρου ξεκινώντας με τον πυρήνα και καταλήγοντας με την διαίρεση του κυτταροπλάσματος από όπου προκύπτουν τελικά τα δύο θυγατρικά κύτταρα. Τέλος ένας οργανισμός έχει την δυνατότητα να «ξεφορτώνεται» ανεπιθύμητα κύτταρα μέσω μιας διαδικασίας προγραμματισμένου κυτταρικού θανάτου που ονομάζεται απόπτωση (apoptosis). Κύτταρα τα οποία δεν είναι πλέον χρήσιμα στον οργανισμό ή κύτταρα τα οποία παύουν να είναι λειτουργικά λόγω πιθανών βλαβών που δεν μπορούν να επιδιορθωθούν οδηγούνται, μέσω ειδικών πρωτεϊνών που ονομάζονται κασπάσες (caspases), σε μια μορφή «οικειοθελούς αυτοκτονίας» την απόπτωση. Μάλιστα σε αντίθεση με την κυτταρική νέκρωση (necrosis) που επέρχεται από κάποιον τραυματισμό και κατά την οποία το κύτταρο καταστρέφεται αδειάζοντας το περιεχόμενό του στα γειτονικά κύτταρα και προκαλώντας πιθανή επιβλαβή για αυτά τα κύτταρα φλεγμονή, η απόπτωση είναι «καθαρή» περίπτωση θανάτου που δεν θέτει σε κίνδυνο τα γειτονικά κύτταρα [4].



Εικόνα 1.5: Ο κυτταρικός κύκλος

1.4 Επαγωγή βλαβών DNA

1.4.1 Οξειδωτικό στρες

Το οξειδωτικό στρες προκαλείται στα κύτταρα από τις λεγόμενες Δραστικές Μορφές Οξυγόνου (Reactive Oxygen Species – ROS). Αυτές είναι κάποιες μορφές μορίων που περιέχουν οξυγόνο και συγκεκριμένα είναι το ιονισμένο οξυγόνο O_2^- , το υπεροξειδίο του υδρογόνου H_2O_2 , οι ρίζες υδροξυλίου OH^- και το μονατομικό ιόν οξυγόνου O^{2-} . Τα συγκεκριμένα μόρια είναι πολύ δραστικά, δηλαδή αντιδρούν εύκολα και γρήγορα με άλλα μόρια προκαλώντας συνήθως την οξείδωση τους και μπορούν να προκαλέσουν βλάβες σε μακρομόρια όπως είναι το DNA, οι πρωτεΐνες και τα λιπίδια. Οι Δραστικές Μορφές Οξυγόνου παράγονται συνεχώς στα κύτταρα ως προϊόν του μεταβολισμού και άλλων βιοχημικών διαδικασιών αλλά και από εξωγενείς παράγοντες όπως είναι το κάπνισμα, διάφορα χημικά και ιοντίζουσες ακτινοβολίες. Πολλοί αμυντικοί μηχανισμοί έχουν αναπτυχθεί στον οργανισμό για να περιορίζουν τα επίπεδα των Δραστικών Μορφών Οξυγόνου και των βλαβών που αυτές προκαλούν [5]. Το οξειδωτικό στρες προκαλείται όταν η παραγωγή των Δραστικών Μορφών Οξυγόνου στον οργανισμό υπερβαίνει τους φυσικούς αντιοξειδωτικούς μηχανισμούς με αποτέλεσμα την συσσώρευση των συγκεκριμένων μορίων. Αυτά μπορούν να προκαλέσουν αρκετές βλάβες στο DNA με κυριότερες την οξείδωση αζωτούχων βάσεων αλλά και τις μονόκλωνες ή δίκλωνες θραύσεις της αλυσίδας του DNA (βλέπε κεφάλαιο 2). Οι Δραστικές Μορφές Οξυγόνου είναι ο πιο συχνός παράγοντας επαγωγής βλαβών στο γενετικό υλικό [6].

1.4.2 Ακτινοβολίες

Είναι γνωστό ότι οι ακτινοβολίες και συγκεκριμένα αυτές που ανήκουν στο κομμάτι εκείνο του ηλεκτρομαγνητικού φάσματος που προκαλούν ιονισμό ή σωματιδιακές ακτινοβολίες με αντίστοιχες ιδιότητες (ιοντίζουσες ακτινοβολίες), μπορούν να προκαλέσουν βλάβες στο DNA των κυττάρων [7]. Οι ιοντίζουσες ακτινοβολίες αναλύονται περιεκτικά στην επόμενη ενότητα της παρούσας

εργασίας. Συνοπτικά αναφέρουμε ότι μπορούν να προκαλέσουν βλάβες στις αζωτούχες βάσεις καθώς και σπάσιμο της μίας ή και των δυο πολυνουκλεοτιδικών αλυσίδων του μορίου του DNA. Χαρακτηριστικό των ιονιζουσών ακτινοβολιών είναι ότι μπορούν να προκαλέσουν τις λεγόμενες σύνθετες βλάβες στο γενετικό υλικό (complex damage), δηλαδή πολλές μεμονωμένες βλάβες συγκεντρωμένες σε μικρό τμήμα του DNA οι οποίες είναι πολύ δύσκολο να επιδιορθωθούν από το κύτταρο και προκαλούν μεταλλάξεις, γονιδιωματική αστάθεια (υψηλή συχνότητα μεταλλάξεων στο γονιδίωμα των κυττάρων) και κυτταρικό θάνατο [8].

1.4.3 Στρες αντιγραφής

Βλάβες στο γενετικό υλικό που παράγονται ενδογενώς από λάθη κατά την αντιγραφή του DNA αναφέρονται ως στρες αντιγραφής και οδηγούν σε καθυστέρηση ή και παύση της διαδικασίας της αντιγραφής. Οι βλάβες αυτές μπορούν να προκύψουν από μια πληθώρα διαφορετικών πηγών που θεωρούνται ότι επηρεάζουν την αντιγραφή όπως λάθη των ενζύμων που πραγματοποιούν την αντιγραφή, τα τελομερή (περιοχές του DNA που δεν φέρουν γενετική πληροφορία αλλά προστατεύουν το υπόλοιπο μόριο από την φθορά), «συγκρούσεις» μεταξύ των μορίων που πραγματοποιούν την μεταγραφή και την μετάφραση, προηγούμενες βλάβες στο DNA κ.α. Αυτοί οι παράγοντες μπορεί να οδηγήσουν σε λάθη στην αντιγραφή με αποτέλεσμα το μόριο DNA που προκύπτει να μην είναι το επιθυμητό ή ακόμα και παύση της αντιγραφής που εμποδίζει την ομαλή λειτουργία του κυτταρικού κύκλου. Οι επιδιορθωτικοί μηχανισμοί για το γενετικό υλικό που έχουν αναπτύξει τα κύτταρα βοηθούν στην διατήρηση της ακεραιότητας του DNA κατά την αντιγραφή όμως κάποια δυσλειτουργία αυτών των μηχανισμών οδηγεί σε γονιδιωματική αστάθεια και πιθανές ασθένειες όπως ο καρκίνος [9].

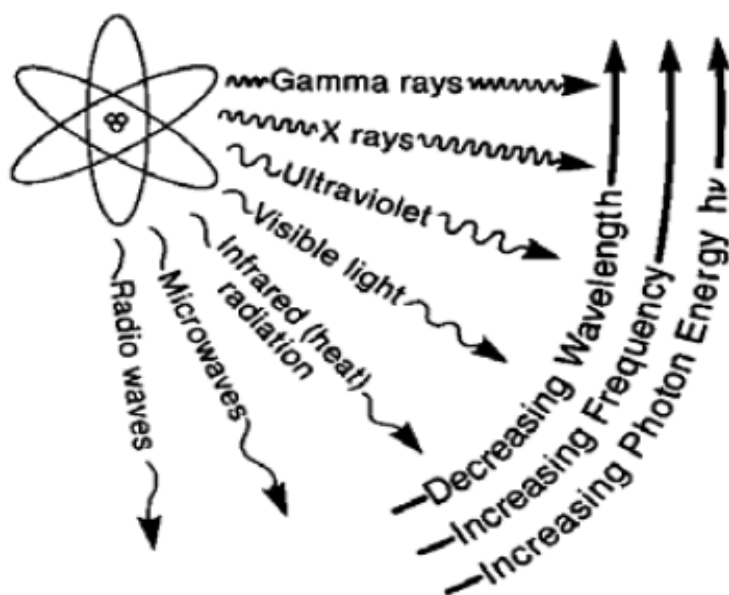
1.5 Ιοντίζουσες ακτινοβολίες

1.5.1 Εισαγωγή

Με τον όρο ακτινοβολία εννοούμε την εκπομπή και μετάδοση ενέργειας, στο κενό ή μέσα σε κάποιο υλικό μέσο, με την μορφή κυμάτων (συνήθως ηλεκτρομαγνητικών) και σωματιδίων. Η ακτινοβολία χωρίζεται σε ιοντίζουσα και μη-ιοντίζουσα ανάλογα με την ενέργεια που φέρουν τα σωματίδια που την αποτελούν. Το βασικό χαρακτηριστικό της ιοντίζουσας ακτινοβολίας είναι ότι φέρει ενέργεια που επαρκεί για τον ιονισμό ατόμων και μορίων, δηλαδή την απομάκρυνση ηλεκτρονίων και την μετατροπή αυτών σε θετικά φορτισμένα ιόντα, καθώς και το σπάσιμο χημικών δεσμών. Οι άνθρωποι εκτίθενται καθημερινά σε πηγές ιοντίζουσας ακτινοβολίας τόσο φυσικές όσο και τεχνητές. Η φυσική ιοντίζουσα ακτινοβολία προέρχεται από διάφορες πηγές όπως τα τουλάχιστον 60 ραδιενεργά ισότοπα που βρίσκονται στο χώμα του εδάφους, το νερό και τον αέρα αλλά και την κοσμική ακτινοβολία που αυξάνεται όσο ανεβαίνουμε υψόμετρο. Υπολογίζεται ότι περίπου το 80% της ετήσιας δόσης ιονιζουσών ακτινοβολιών που δέχεται ένα άτομο προέρχεται από φυσικές πηγές. Τεχνητές πηγές ακτινοβολίας αποτελούν οι σταθμοί πυρηνικής ενέργειας και οι ιατρικές συσκευές που χρησιμοποιούνται για διάγνωση και θεραπεία όπως για παράδειγμα οι αξονικοί τομογράφοι και η ακτινοθεραπεία [10]. Χαρακτηριστικό των ιονιζουσών ακτινοβολιών και που θα αναλύσουμε παρακάτω είναι ότι προκαλούν τις λεγόμενες σύνθετες ή ομαδοποιημένες βλάβες στο DNA των κυττάρων, κάτι το οποίο μπορεί να ιδωθεί σαν θετικό όταν οδηγεί σε θανάτωση καρκινικών κυττάρων κατά την ακτινοθεραπεία αλλά και σαν αρνητικό όταν οδηγεί στην δημιουργία μεταλλάξεων και γονιδιωματικής αστάθειας σε φυσιολογικούς ιστούς [11].

1.5.2 Τυπικές ιοντίζουσες ακτινοβολίες

Η ιοντίζουσα ακτινοβολία μπορεί να χωριστεί στην ηλεκτρομαγνητική ιοντίζουσα ακτινοβολία και την σωματιδιακή ιοντίζουσα ακτινοβολία. Όπως είπαμε χαρακτηριστικό των ιοντίζουσών ακτινοβολιών είναι ότι φέρουν ενέργεια ικανή να προκαλέσει ιονισμό στα άτομα του μέσου στο οποίο διαδίδεται [10]. Από ολόκληρο το φάσμα της ηλεκτρομαγνητικής ακτινοβολίας (βλέπε εικόνα 2.1) το τμήμα με τις υψηλότερες συχνότητες και άρα και τις υψηλότερες ενέργειες είναι εκείνο που μπορεί να χαρακτηριστεί ως ιοντίζουσα ακτινοβολία δηλαδή οι ακτίνες-Χ, οι ακτίνες-γ και η «βαθιά» υπεριώδης ακτινοβολία (UVC). Για τις σωματιδιακές ιοντίζουσες ακτινοβολίες τυπικά παραδείγματα αποτελούν τα σωματίδια α και β, ακτίνες πρωτονίων, νετρονίων αλλά και φορτισμένα ιόντα. Τα σωματίδια αυτά είτε προκύπτουν ως προϊόντα ραδιενεργών διασπάσεων είτε επιταχύνονται τεχνητά από ειδικούς επιταχυντές ώστε να αποκτήσουν την απαραίτητη ενέργεια [12]. Σημαντικό είναι και το γεγονός ότι τα σωματίδια προκαλούν άμεσους ιονισμούς ενώ αντίθετα η δράση των φωτονίων (ηλεκτρομαγνητική ακτινοβολία) λέμε ότι είναι έμμεση αφού δεν προκαλούν απευθείας ιονισμό στα άτομα του υλικού-στόχου.



Εικόνα 2.1: Το φάσμα της ηλεκτρομαγνητικής ακτινοβολίας.

1.5.3 Στάδια αλληλεπίδρασης ακτινοβολίας και βιολογικής ύλης

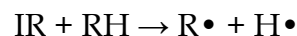
Όταν μια ιοντίζουσα ακτινοβολία προσπέσει σε κάποιο βιολογικό υλικό - στόχο, όπως ένας ανθρώπινος ιστός, αλληλεπιδρά μαζί του σε συγκεκριμένα στάδια [13]. Ανεξάρτητα από το είδος της ιοντίζουσας ακτινοβολίας, αυτή θα οδηγήσει σε απορρόφηση ενέργειας μέσω ιονισμών και διεγέρσεων προκαλώντας βλάβες στο γενετικό υλικό, σπάσιμο χημικών δεσμών και δημιουργώντας δραστικές μορφές οξυγόνου ενώ θα επιφέρει και κάποια βιολογικά αποτελέσματα είτε αυτό είναι ο κυτταρικός θάνατος, κάποια μετάλλαξη κλπ. Όμως στο ενδιάμεσο μεσολαβούν και άλλα στάδια δράσης της ακτινοβολίας ενώ και η χρονική διάρκεια κάθε σταδίου αλληλεπίδρασης μπορεί να διαφέρει σημαντικά από τα προηγούμενα. Έτσι διακρίνουμε τα εξής στάδια αλληλεπίδρασης από την στιγμή της πρόσπτωσης της ακτινοβολίας στο υλικό - στόχο [13]:

- Φυσικό στάδιο: Η χρονική διάρκεια είναι της τάξης των asec μέχρι fsec (10^{-18} έως 10^{-15} δευτερόλεπτα) και είναι το στάδιο της εναπόθεσης ενέργειας μέσω των ιονισμών και των διεγέρσεων, όπως αναφέραμε.
- Φυσικοχημικό στάδιο: Διαρκεί από μερικά fsec έως και μερικά nsec (10^{-15} έως 10^{-9} δευτερόλεπτα) και τότε έχουμε το σπάσιμο χημικών δεσμών, την ραδιόλυση του ύδατος και τον σχηματισμό δραστικών μορφών οξυγόνου (άμεση και έμμεση δράση της ακτινοβολίας - βλέπε επόμενη ενότητα).
- Χημικό στάδιο: Η χρονική κλίμακα αυτού του σταδίου είναι από δισεκατομμυριοστά μέχρι χιλιοστά του δευτερολέπτου (nsec έως msec) και σε αυτή την φάση οι δραστικές μορφές οξυγόνου αντιδρούν είτε μεταξύ τους, είτε με τα μόρια που τις περιβάλλουν και δημιουργούνται οι βλάβες στα βιολογικά μακρομόρια.
- Βιοχημικό στάδιο: Διαρκεί από χιλιοστά του δευτερολέπτου έως και μερικές ώρες και είναι το στάδιο στο οποίο το κύτταρο αναγνωρίζει την ύπαρξη βλαβών στο γενετικό του υλικό, η σύνθεση πρωτεϊνών αναστέλλεται και ξεκινάει η λειτουργία των μηχανισμών επιδιόρθωσης.

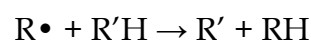
- Κυτταρικό στάδιο: Είναι το τελευταίο στάδιο αλληλεπίδρασης της ακτινοβολίας με την βιολογική ύλη στο οποίο και παρατηρούνται τα αποτελέσματα της ακτινοβόλησης. Η διάρκεια του είναι από κάποιες ώρες μέχρι και μήνες μετά την πρόσπτωση της ακτινοβολίας. Τα φαινόμενα που παρατηρούνται σε αυτό το στάδιο έρχονται ως αποτέλεσμα της πρόκλησης βλαβών (συνήθως σύνθετων) στο DNA του κυττάρου τις οποίες το κύτταρο αδυνατεί να επιδιορθώσει. Μπορούμε να έχουμε αναστολή της κυτταρικής διαίρεσης και αδρανοποίηση του κυττάρου, κυτταρικό θάνατο, δημιουργία μεταλλάξεων στο γενετικό υλικό, βλάβη σε ιστούς κλπ. Με τα χρόνια μπορούν να επέλθουν διάφορες βιολογικές αντιδράσεις στον οργανισμό όπως ανοσολογικές αντιδράσεις, λειτουργικές βλάβες, καρκινογένεση κ.α.

1.5.4 Άμεσα και έμμεσα αποτελέσματα δράσης της ακτινοβολίας

Μια ιοντιζουσα ακτινοβολία που προσπίπτει σε έναν βιολογικό στόχο είναι γνωστό ότι μπορεί να προκαλέσει βλάβες σε βιομόρια με δύο τρόπους, είτε άμεσα είτε έμμεσα (εικόνα 2.2)[12]. Στην άμεση δράση της ακτινοβολίας έχουμε την απευθείας πρόκληση βλαβών σε μόρια του κυττάρου με βασικό το DNA αλλά και πρωτεΐνες, ένζυμα και RNA. Ένας ή και παραπάνω χημικοί δεσμοί σπάνε και απομένουν άτομα ή μόρια με ασύζευκτα ηλεκτρόνια τα οποία είναι πολύ δραστικά και αντιδρούν με τα μόρια που βρίσκονται γύρω τους. Η άμεση δράση της ακτινοβολίας απεικονίζεται στην παρακάτω αντίδραση όπου με RH συμβολίζεται κάποιο τυχαίο βιομόριο και με IR (ionizing radiation) μια ιοντιζουσα ακτινοβολία:

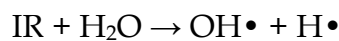


Ενώ και τα δυο προϊόντα της αντίδρασης μπορούν να αντιδράσουν με άλλα βιομόρια (R'H) όπως DNA, λιπίδια και πρωτεΐνες, για παράδειγμα σύμφωνα με την αντίδραση

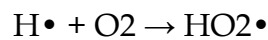


Τα μόρια που δημιουργούνται έχουν διαφορετικές χημικές ιδιότητες από τα αρχικά.

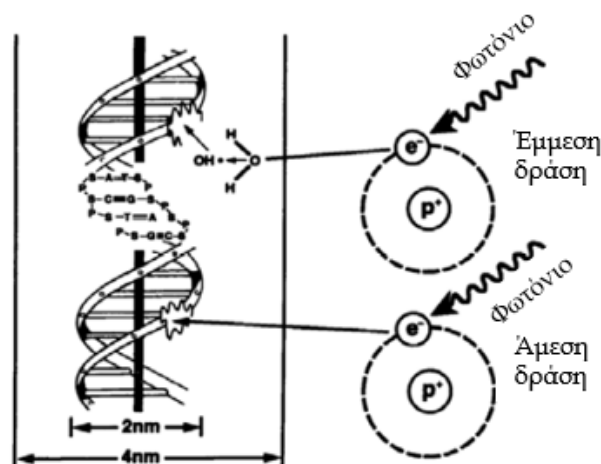
Η έμμεση δράση της ακτινοβολίας γίνεται με την λεγόμενη ραδιόλυση του ύδατος. Το νερό αποτελεί το 80% της μάζας του κυττάρου και έτσι είναι πολύ πιθανό η ιοντιζουσα ακτινοβολία να χτυπήσει κάποιο μόριο νερού διασχίζοντας το κύτταρο. Συγκεκριμένα υπολογίζεται ότι περίπου τα δυο τρίτα των προκαλούμενων βλαβών από ακτινοβολίες χαμηλής LET οφείλονται στον έμμεσο μηχανισμό [14]. Η αντίδραση της ραδιόλυσης του ύδατος απεικονίζεται ως:



Τα προϊόντα αυτά είναι ιδιαίτερα δραστικά. Οι ρίζες υδροξυλίου $\text{OH}\cdot$ ανήκουν στις Δραστικές Μορφές Οξυγόνου οι οποίες όπως έχουμε ήδη αναφέρει είναι ιδιαίτερα επικίνδυνες για το κύτταρο γιατί προκαλούν οξειδώσεις σε βιομόρια. Τα άτομα υδρογόνου μπορούν να αντιδράσουν με μόρια οξυγόνου σύμφωνα με την αντίδραση:



Όπου και το υπεροξείδιο του υδρογόνου το έχουμε δει στις Δραστικές Μορφές Οξυγόνου. Αντιλαμβανόμαστε επομένως ότι μέσω της αλληλεπίδρασης με την ιοντιζουσα ακτινοβολία το νερό στο εσωτερικό του κυττάρου γίνεται ένας επικίνδυνος παράγοντας για την επιβίωση του κυττάρου αφού τα παράγωγα της αντίδρασης απειλούν άμεσα τα βιομόρια [14]



Εικόνα 2.2: Η άμεση και έμμεση δράση της ακτινοβολίας.

Κεφάλαιο 2: Βλάβες DNA

2.1 Η Γραμμική Εναπόθεση Ενέργειας και πυκνότητα βλαβών

Μια σημαντική έννοια για τις ιοντίζουσες ακτινοβολίες είναι αυτή της Γραμμικής Εναπόθεσης Ενέργειας (Linear Energy Transfer - LET). Όταν μια ιοντίζουσα ακτινοβολία διασχίζει την ύλη χάνει σταδιακά ενέργεια λόγω των αλληλεπιδράσεων με τα άτομα του υλικού-στόχου. Για ένα συγκεκριμένο υλικό-στόχο, το ποσοστό της ενέργειας που χάνεται εξαρτάται από τον τύπο και την ενέργεια της ακτινοβολίας και από την πυκνότητα του υλικού-στόχου [12]. Η εναποτιθέμενη ενέργεια σε ένα υλικό όπως για παράδειγμα ο ανθρώπινος ιστός, ονομάζεται Γραμμική Εναπόθεση Ενέργειας και ορίζεται ως η μέση ενέργεια που εναποτίθεται ανά μονάδα μήκους της διαδρομής της ακτινοβολίας ενώ οι μονάδες μέτρησής της είναι τα keV/μm (κιλοηλεκτρονιοβόλτ ανά εκατομμυριοστό του μέτρου). Είναι σημαντικό ότι η Γραμμική Εναπόθεση Ενέργειας μεταβάλλεται κατά μήκος της διαδρομής επειδή καθώς τα σωματίδια χάνουν την ενέργεια τους λόγω των αλληλεπιδράσεων, επιβραδύνονται και αντιδρούν ακόμα πιο εύκολα με τα άτομα που βρίσκονται γύρω τους. Έτσι το ποσοστό της μεταφερόμενης ενέργειας αυξάνεται κατά μήκος της διαδρομής και παρατηρείται μια κορυφή της εναπόθεσης ενέργειας η λεγόμενη «κορυφή Bragg». Η Γραμμική Εναπόθεση Ενέργειας είναι σημαντική επειδή η βιολογική επίδραση της ακτινοβολίας, δηλαδή η πρόκληση ιονισμών και διεγέρσεων σε βιολογικά μόρια και κυρίως στο DNA των κυττάρων, εξαρτάται άμεσα από την μέση τιμή της πρώτης. Οι ακτινοβολίες κατηγοριοποιούνται σε χαμηλής και υψηλής LET, με τις σωματιδιακές ακτινοβολίες να έχουν συνήθως υψηλές τιμές ενώ αντίθετα οι ακτίνες γ και Χ είναι χαμηλής LET. Γενικά η βιολογική δράση της ακτινοβολίας αυξάνεται όσο αυξάνεται η LET μέχρι μια οριακή τιμή περίπου 100 keV/μm η οποία όταν ξεπεραστεί παρατηρείται μείωση της βιολογικής δράσης. Για τις ακτινοβολίες χαμηλής LET, δηλαδή ακτινοβολίες φωτονίων, οι ιονισμοί που

προκαλούν στο γενετικό υλικό είναι αραιοί μεταξύ τους στα άτομα και μόρια του στόχου αφού οι διαστάσεις των φωτονίων είναι μικρότερες σε σύγκριση με τις διαστάσεις των βιομορίων όπως το DNA. Αντίθετα με τις σωματιδιακές ακτινοβολίες παρατηρούνται πιο πυκνοί ιονισμοί ακριβώς λόγω του ότι οι διαστάσεις των σωματιδίων αυτών είναι συγκρίσιμες με αυτές των βιομορίων [12].

2.2 Είδη βλαβών DNA

Σε αυτή την ενότητα θα αναλύσουμε τις βλάβες (εικόνα 2.3) που προκαλούνται στο γενετικό υλικό των κυττάρων τόσο κατά την ακτινοβολή τους με ιοντίζουσες ακτινοβολίες αλλά και από άλλους παράγοντες όπως το στρες αντιγραφής και το οξειδωτικό στρες (βλέπε και ενότητα 1.4). Οι βλάβες στο DNA μπορούν να χωριστούν σε μεμονωμένες/ απλές και ομαδοποιημένες/ σύνθετες. Οι μεμονωμένες βλάβες είναι αρκετά συνηθισμένες στο γενετικό υλικό, πολλές φορές δημιουργούνται από ενδογενείς παράγοντες και συνήθως επιδιορθώνονται με επιτυχία από το κύτταρο. Αντίθετα οι ομαδοποιημένες βλάβες είναι χαρακτηριστικό κυρίως εξωγενών παραγόντων όπως οι ακτινοβολίες και σε σπάνιες περιπτώσεις και για συγκεκριμένους τύπους κυττάρων μπορεί να δημιουργηθούν ενδογενώς [15].

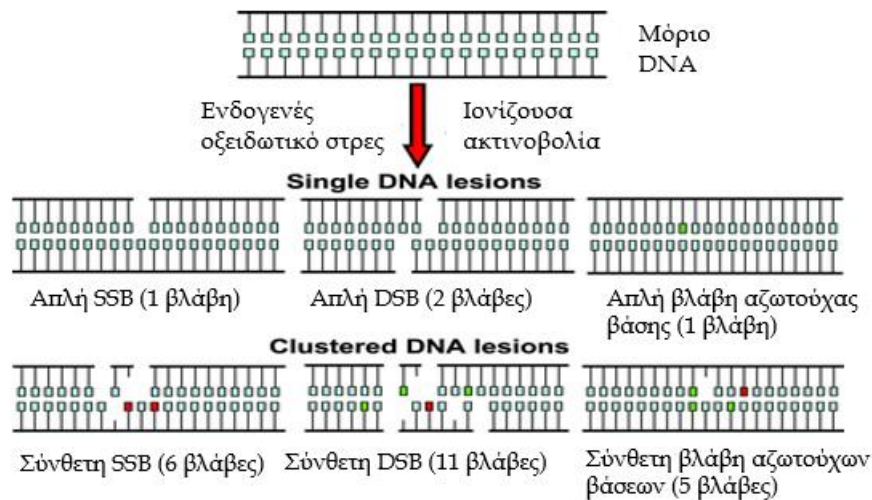
2.2.1 Μεμονωμένες βλάβες

Οι κυριότερες βλάβες που μπορούν να προκύψουν στο γενετικό υλικό είναι μονόκλωνες θραύσεις (Single strand break – SSB) δηλαδή σπάσιμο της μιας από τις δύο αλυσίδες του μορίου του DNA, δίκλωνες θραύσεις (Double strand break – DSB) δηλαδή σπάσιμο και των δύο αλυσίδων, βλάβη σε κάποια αζωτούχα βάση (Base damage/lesion) και δια-δεσμοί (crosslinks) μεταξύ DNA και πρωτεϊνών όπως οι ιστόνες [12]. Συγκεκριμένα οι αζωτούχες βάσεις μπορούν να οξειδωθούν εξαιτίας για παράδειγμα των Δραστικών Μορφών Οξυγόνου, να υποστούν μεθυλίωση ή ακόμα και να καταστραφούν εντελώς αν σπάσει ο δεσμός μεταξύ της δεοξυριβόζης και της αζωτούχας βάσης, δημιουργώντας τα λεγόμενα αβασικά

(abasic) σημεία [16]. Σε περίπτωση που η βλάβη προκληθεί σε κάποιο σημείο της ραχοκοκαλιάς σακχάρου – φωσφορικής ομάδας όπως για παράδειγμα σε κάποιον από τους άνθρακες της δεοξυριβόζης μπορεί να σπάσει ο δεσμός ανάμεσα σε δύο νουκλεοτίδια και να έχουμε θραύση του ενός ή και των δύο αλυσίδων στο DNA. Οι μεμονωμένες βλάβες αντιμετωπίζονται με τους κατάλληλους επιδιορθωτικούς μηχανισμούς από το κύτταρο και τις περισσότερες φορές δεν απειλούν την επιβίωση του κυττάρου.

2.2.2 Ομαδοποιημένες βλάβες

Οι ομαδοποιημένες βλάβες, τις οποίες θα αναλύσουμε εκτενώς στην συνέχεια, προκύπτουν όταν έχουμε δύο ή και παραπάνω βλάβες σε μικρή απόσταση μεταξύ τους (συνήθως αναφέρεται απόσταση μικρότερη των δέκα ζευγών βάσεων) στο μόριο του DNA [17]. Δηλαδή όταν πολλές από τις μεμονωμένες βλάβες που αναφέραμε συγκεντρώνονται σε πολύ κοντινά σημεία πάνω στο μόριο και δημιουργούν τα clusters όπως αναφέρονται συχνά. Οι ομαδοποιημένες βλάβες στο γενετικό υλικό αποτελούν την «υπογραφή» των ιονιζουσών ακτινοβολιών και είναι γνωστό ότι ο αριθμός των βλαβών και η πολυπλοκότητα τους αυξάνεται όσο αυξάνεται η Γραμμική Εναπόθεση Ενέργειας της ακτινοβολίας [18]. Οι πολύπλοκες βλάβες αποτελούν μεγάλο πρόβλημα για το κύτταρο και την επιβίωση του, αφού η επιδιόρθωση τους είναι εξαιρετικά δύσκολη. Οι κυτταρικοί μηχανισμοί επιδιόρθωσης είναι σχεδιασμένοι για να παίρνουν θέση και να διορθώνουν τις εκάστοτε βλάβες ξεχωριστά, όταν όμως αυτές συσσωρεύονται χάνουν σε σημαντικό ποσοστό την αποτελεσματικότητά τους, τόσο στην προσδεσιμότητα στην περιοχή όσο και στη λειτουργία τους. Η επιδιόρθωση γίνεται αργή και πολλές φορές αναποτελεσματική και το κύτταρο οδηγείται σε γονιδιωματική αστάθεια, γήρανση ή και απόπτωση.



Εικόνα 2.3: Μεμονωμένες (single) και ομαδοποιημένες (clustered) βλάβες στο DNA. Παρατηρούνται σπάσιμο της μιας αλυσίδας (SSB), σπάσιμο και των δύο αλυσίδων (DSB) και βλάβη σε βάσεις.

2.3 Επιδιορθωτικοί μηχανισμοί

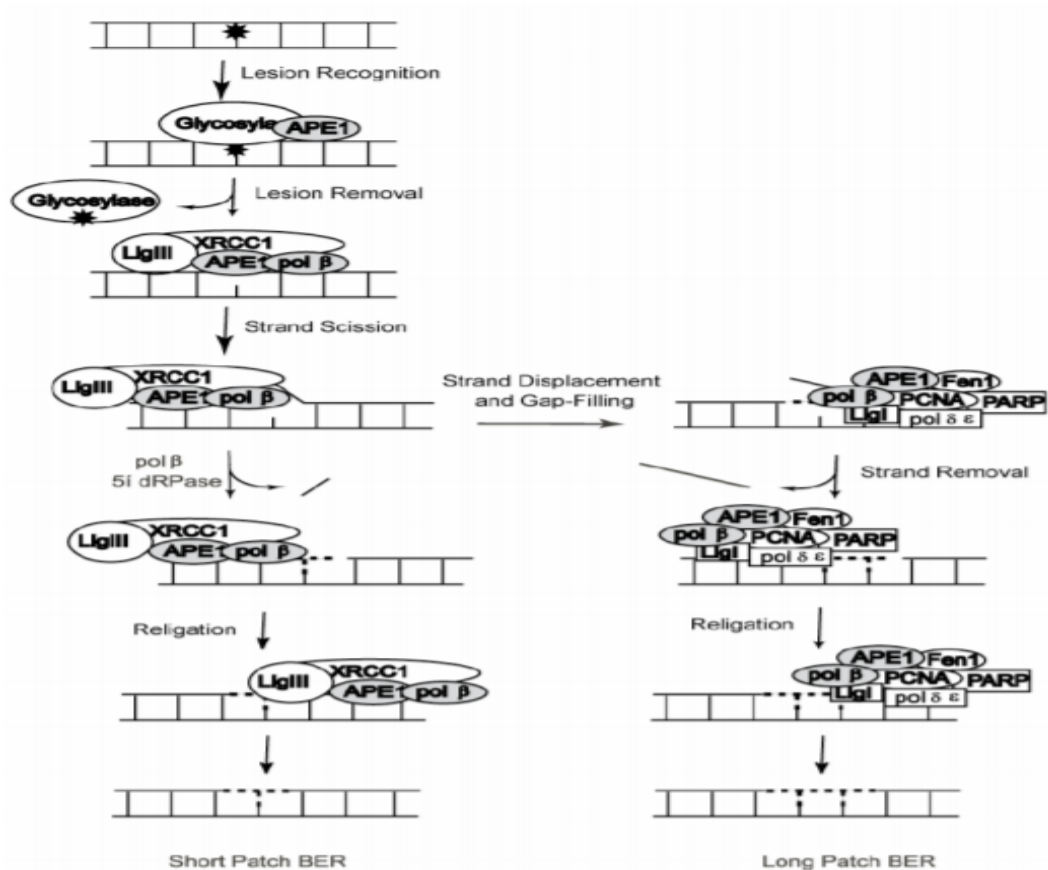
2.3.1 Εισαγωγή

Στο κεφάλαιο αυτό θα περιγράψουμε αναλυτικά τους μηχανισμούς τους οποίους χρησιμοποιεί το ανθρώπινο κύτταρο για να επιδιορθώσει τις βλάβες στο DNA. Υπάρχουν πέντε γνωστοί μηχανισμοί τους οποίους επιστρατεύει το κύτταρο όταν εντοπίσει κάποια βλάβη στο γενετικό του υλικό, ανάλογα τον τύπο της βλάβης, ενώ μπορούν να χρησιμοποιηθούν και σε συνδυασμό μεταξύ τους. Όταν υπάρχει βλάβη μόνο στην μια από τις δύο συμπληρωματικές αλυσίδες του DNA (single-strand damage) το κύτταρο χρησιμοποιεί τρεις μηχανισμούς εξαγωγής-επιδιόρθωσης: τον Μηχανισμό Εκτομής Βάσης (Base Excision Repair), τον Μηχανισμό Εκτομής Νουκλεοτιδίων (Nucleotide Excision Repair) και την Επιδιόρθωση Αναντιστοιχίας (Mismatch Repair). Όταν υπάρχει βλάβη και στις δύο αλυσίδες του DNA (double-strand breaks) το κύτταρο επιστρατεύει τους άλλους δύο επιδιορθωτικούς μηχανισμούς: τον Ομόλογο Ανασυνδυασμό (Homologous Recombination) και την Μη-ομόλογη Ένωση Ακρων (Non-homologous End Joining) [19].

2.3.2 Μηχανισμός Εκτομής Βάσης (BER)

Νουκλεοβάσεις με μικρές χημικές μεταβολές, οι οποίες δεν επηρεάζουν σε μεγάλο βαθμό την δομή της διπλής έλικας του DNA, αποτελούν αντικείμενο επεξεργασίας και επιδιόρθωσης του Μηχανισμού Εκτομής Βάσης (Base Excision Repair – BER) [20]. Ο BER είναι ο κυρίαρχος μηχανισμός για την επιδιόρθωση μικρών βλαβών οι οποίες μπορεί να προκύψουν τόσο από εξωγενείς όσο και από ενδογενείς παράγοντες. Γνωρίζουμε ότι καθημερινά προκύπτουν περίπου δέκα χιλιάδες τέτοιες βλάβες ανά ανθρώπινο κύτταρο επομένως η σημασία του μηχανισμού αυτού είναι τεράστια. Ο μηχανισμός χωρίζεται στον «μικρής έκτασης-BER» (short-patch), στον οποίο μόνο ένα νουκλεοτίδιο επιδιορθώνεται και ο οποίος αποτελεί το μεγαλύτερο ποσοστό δράσης του συγκεκριμένου μηχανισμού και στον «μεγάλης έκτασης-BER» (long-patch), ο οποίος επιδιορθώνει από 2 μέχρι 13 ελαττωματικά νουκλεοτίδια.

Η δράση του BER (σχήμα 3.1) ξεκινάει με μια γλυκοσυλάση η οποία αναγνωρίζει και απομακρύνει την ελαττωματική βάση. Στην συνέχεια η AP ενδονουκλεάση (APE1) καταλύει, μέσω της υδρόλυσης το κόψιμο της αλυσίδας σε εκείνο το αβασικό σημείο. Από αυτό το σημείο ο BER διαχωρίζεται στους δύο επιμέρους μηχανισμούς. Στην περίπτωση του «μικρής έκτασης-BER» η DNA πολυμεράση-β (polβ) γεμίζει το κενό με την σωστή βάση προετοιμάζοντας την ένωση της αλυσίδας, η οποία πραγματοποιείται από ένα σύμπλοκο της DNA λιγκάσης IIIα (LigIIIα) με την XRCC1. Στον «μεγάλης έκτασης-BER», η δράση της polβ ενισχύεται από την πολυμεράση δ ή ε (polδ/ε), την PCNA (proliferating cell nuclear antigen) και διάφορες άλλες πρωτεΐνες όπως η Fen1 και η PARP1 για να συνθέσουν το DNA. Η ένωση της αλυσίδας πραγματοποιείται από την λιγκάση I (LigI) [21].



Σχήμα 3.1: Ο μηχανισμός BER

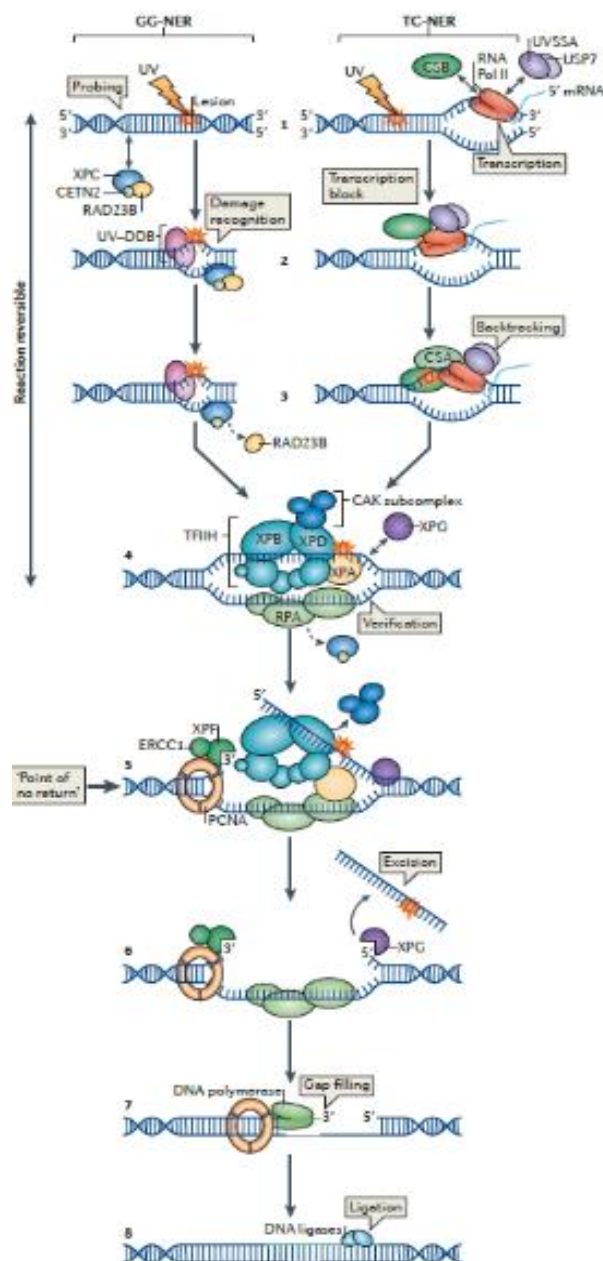
2.3.3 Μηχανισμός Εκτομής Νουκλεοτιδίων (NER)

Ο μηχανισμός εκτομής νουκλεοτιδίων είναι το βασικό επιδιορθωτικό «μονοπάτι» (pathway) που χρησιμοποιείται από τους οργανισμούς για την επιδιόρθωση σοβαρών και εκτεταμένων βλαβών στο γενετικό υλικό, οι οποίες μπορεί να έχουν αρνητική επίδραση στην δομή της διπλής έλικας του DNA, όπως αυτές που προκαλούνται από την υπερϊώδη ακτινοβολία και άλλους περιβαλλοντικούς μεταλλαξογόνους παράγοντες καθώς και από ορισμένους αντικαρκινικούς παράγοντες [22]. Σε αντίθεση με τον μηχανισμό εκτομής βάσης, ο οποίος αντικαθιστά την ελαττωματική αζωτούχα βάση από ένα νουκλεοτίδιο, ο

μηχανισμός εκτομής νουκλεοτιδίων αντικαθιστά ολόκληρα τα νουκλεοτίδια και συνεπώς είναι πιο σύνθετος και περιλαμβάνει την δράση περισσότερων από 25 πρωτεϊνών. Ο NER χωρίζεται σε δύο υποκατηγορίες: Τον NER ολικού γονιδιώματος (Global-Genome NER) και τον NER-μεταγραφής (Transcription-coupled NER). Ο GG-NER ελέγχει ολόκληρο το γονιδίωμα για αλλοιώσεις στην αλυσίδα του DNA, ενώ ο TC-NER ενεργοποιείται όταν η RNA πολυμεράση «κολλήσει» σε κάποιο σημείο της μεταγραφής εξαιτίας βλάβης στην αλυσίδα που χρησιμοποιείται για την μεταγραφή [23].

Στον GG-NER (σχήμα 3.2) ένα σύμπλοκο πρωτεϊνών που αποτελείται από τις XPC, RAD23B (UV excision repair protein RAD23 homologue B) και CETN2 (centrin2) ερευνά συνεχώς το γενετικό υλικό για τον εντοπισμό βλαβών που διαταράσσουν την δομή της διπλής έλικας (βήμα 1 αριστερή στήλη σχήμα 3.2), τις οποίες αναγνωρίζει με την βοήθεια και της UV-DDB (ultra-violet radiation DNA damage-binding protein)(βήμα 2). Μόλις το σύμπλοκο προσδεθεί στο ελαττωματικό σημείο, η RAD23B αποσυνδέεται και απομακρύνεται (βήμα 3). Στον NER-μεταγραφής, η αναγνώριση των βλαβών γίνεται με έμμεσο τρόπο, κατά την μεταγραφή του DNA σε mRNA από την RNA πολυμεράση II (RNA pol II). Όσο δρα η RNA pol II για την μεταγραφή του DNA, αλληλεπιδρά με τις UVSSA (UV-stimulated scaffold protein A), USP7 (ubiquitin-specific-processing protease 7) και CSB (Cockayne syndrome protein)(βήμα 1 δεξιά στήλη σχήμα 3.2). Μόλις η RNA pol II μπλοκάρει σε κάποια βλάβη στην αλυσίδα του DNA, η σύνδεση της CSB με την πολυμεράση αυξάνεται (βήμα 2) και δημιουργείται το σύμπλοκο CSA-CSB το οποίο έχει ως αποτέλεσμα την αντιστροφή της πορείας (backtracking) της πολυμεράσης για να γίνει το ελαττωματικό σημείο προσβάσιμο για επιδιόρθωση (βήμα 3). Μετά την αναγνώριση της βλάβης το σύμπλοκο TFIIH (transcription initiation factor III) προσέρχεται στο σημείο και στις δύο περιπτώσεις του NER, ενώ επίσης είτε μαζί με το TFIIH, είτε χωριστά προσέρχεται και η XPG. Μετά την πρόσδεση του TFIIH στο σημείο το υπο-σύμπλοκο CAK (το οποίο αποτελεί μέρος του TFIIH) απομακρύνεται (βήμα 4). Το σύμπλοκο TFIIH ανοίγει περαιτέρω την αλυσίδα του DNA, ενώ οι XPD, XPB και XPA, η οποία προσδένεται σε νουκλεοτίδια που παρουσιάζουν χημικές μεταβολές, επιβεβαιώνουν την ύπαρξη βλάβης. Επίσης σε

αυτό το στάδιο η RPA (replication protein A) προσέρχεται για να καλύψει και να προστατεύσει την μη-ελαττωματική αλυσίδα. Στην συνέχεια το διμερές XPF-ERCC1 κόβει την ελαττωματική αλυσίδα (βήμα 5), ενώ από την άλλη μεριά η XPG ολοκληρώνει την τομή απομακρύνοντας τελικά ένα τμήμα 22-30 νουκλεοτιδίων (βήμα 6). Η PCNA, η οποία προσδένεται στο σημείο μετά το κόψιμο της αλυσίδας από το διμερές XPF-ERCC1, «στρατολογεί» μια από τις πολυμεράσες DNA Pol δ, DNA Pol κ ή DNA Pol ε για να ξεκινήσει η σύνθεση της νέας αλυσίδας (βήμα 7), ενώ ο μηχανισμός ολοκληρώνει την λειτουργία του με την ένωση της νέας αλυσίδας με την παλιά μέσω της DNA λιγκάσης 1 ή 3 (βήμα 8) [23].



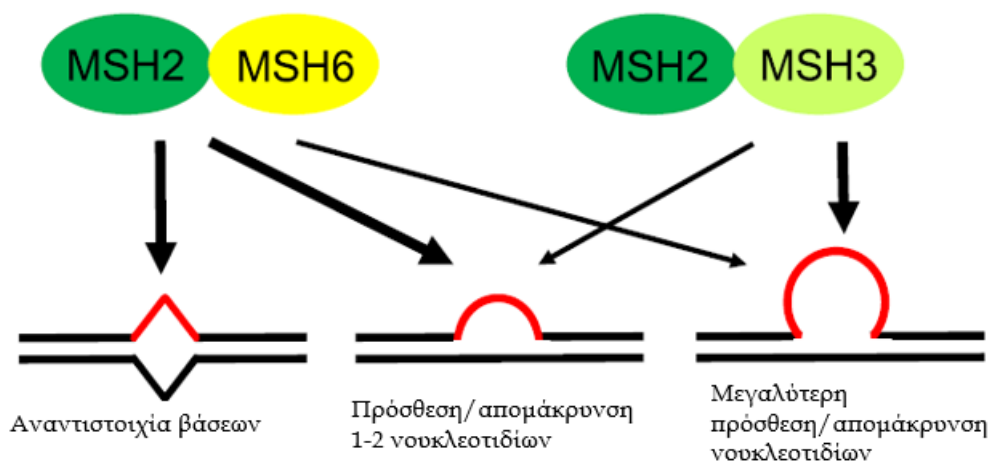
Σχήμα 3.2: Ο μηχανισμός NER

2.3.4 Επιδιόρθωση αναντιστοιχίας (MMR)

Η επιδιόρθωση αναντιστοιχίας (mismatch repair) είναι ένα από τα αρκετά βιολογικά «μονοπάτια» (pathways) το οποίο έχει διατηρηθεί σε μεγάλο βαθμό κατά την διάρκεια της εξέλιξης από τα βακτήρια στους ανθρώπους. Ο αντιστοιχος μηχανισμός του βακτηρίου *Escherichia Coli* έχει μελετηθεί εκτενώς, τόσο βιοχημικά, όσο και γενετικά και αποτελεί πρότυπο για την κατανόηση της επιδιόρθωσης αναντιστοιχίας στους ευκαρυωτικούς οργανισμούς. Ο ρόλος αυτού του μηχανισμού είναι να επιδιορθώνει τις λανθασμένες αντιστοιχίες μεταξύ των αζωτούχων βάσεων αλλά και το τυχόν πλεόνασμα ή έλλειμα νουκλεοτιδίων (τοποθέτηση περιττών βάσεων ή μη τοποθέτηση κάποιας βάσης) που μπορεί να προκύψει κατά την αντιγραφή του DNA, αφαιρώντας την «προβληματική» βάση και τοποθετώντας μια σωστή στην θέση της. Με αυτό τον τρόπο η επιδιόρθωση αναντιστοιχίας προλαμβάνει τις μεταλλάξεις που συμβαίνουν στο γενετικό υλικό (κατά κύριο λόγο αυθόρμητα στην διάρκεια της αντιγραφής) από το να γίνουν μόνιμες μέσω της κυτταρικής διαίρεσης, αυξάνοντας ταυτόχρονα κατά πολύ μεγάλα ποσοστά την ακρίβεια της αντιγραφής του DNA [24].

Πολλές από τις ανθρώπινες πρωτεΐνες που συμμετέχουν στον μηχανισμό MMR έχουν ανακαλυφθεί με βάση τα ομόλογα τους στις αντίστοιχες πρωτεΐνες του βακτηρίου *E. Coli* όπως οι MutS α,β και MutL α,β,γ ή EXO1, RPA, PCNA, DNA polymerase δ και DNA ligase I. Όμως ενώ στο βακτήριο οι MutS και MutL είναι ομο-διμερή τα ανθρώπινα ομόλογα είναι έτερο-διμερή. Συγκεκριμένα η MSH2 συνδέεται με την MSH6 ή με την MSH3 σχηματίζοντας αντίστοιχα τα MutS α και MutS β . Οι δύο αυτές πρωτεΐνες παίζουν βασικό ρόλο στην αναγνώριση αναντιστοιχιών και στην εκκίνηση του επιδιορθωτικού μηχανισμού. Η MutS α αναγνωρίζει και επιδιορθώνει κυρίως αναντιστοιχίες μεταξύ βάσεων και πλεονάσματα/ελλείματα 1-2 νουκλεοτιδίων, ενώ η MutS β επιδιορθώνει κατά κύριο λόγο μεγαλύτερα πλεονάσματα/ελλείματα νουκλεοτιδίων (βλέπε σχήμα 3.3). Αντίστοιχα η MLH1 συνδέεται με τις PMS2, PMS1 ή MLH3 σχηματίζοντας τις MutL α , MutL β και MutL γ αντίστοιχα. Η PCNA συμμετέχει στην έναρξη του μηχανισμού επιδιόρθωσης, ενώ επίσης αλληλεπιδρά με τις MSH2 και MLH1

βοηθώντας στην σύνθεση του DNA μετά την αφαίρεση των ελλατωματικών βάσεων. Στην σύνθεση του DNA χρησιμοποιείται φυσικά και η DNA πολυμεράση δ (Polδ). Η RPA φαίνεται να συμμετέχει σε όλα τα στάδια δράσης του μηχανισμού αφού προσδένεται στο ελλατωματικό σημείο πριν την MutSα, βοηθάει στην αφαίρεση των βάσεων που έχουν τοποθετηθεί λανθασμένα, προστατεύει το μέρος του γενετικού υλικού που δεν έχει υποστεί κάποια βλάβη και διευκολύνει την σύνθεση νέου DNA. Η EXO1 χρησιμεύει και αυτή στην αφαίρεση βάσεων, ενώ τέλος η DNA λιγκάση συνδέει τις δύο αλυσίδες του DNA μετά την δράση του επιδιορθωτικού μηχανισμού [25].



Σχήμα 3.3: Η MutSα και MutSβ στον MMR

2.3.5 Ομόλογος ανασυνδυασμός (HR)

Ο μηχανισμός του ομόλογου ανασυνδυασμού (Homologous Recombination – σχήμα 3.4) χρησιμοποιείται για την επιδιόρθωση χρωμοσωμάτων στα οποία εμφανίζονται δίκλωνες θραύσεις του DNA (double-strand breaks) δηλαδή σπάσιμο και των δύο αλυσίδων της διπλής έλικας σε απόσταση μικρότερη των δέκα ζευγών βάσεων μεταξύ τους. Ο ομόλογος ανασυνδυασμός συναντάται σε όλες τις μορφές ζωής, ενώ μέσω της μελέτης του μηχανισμού στο βακτήριο *Escherichia Coli* και στον μύκητα *Saccharomyces Cerevisiae* έγινε γνωστός ο ρόλος του στην επιδιόρθωση του γενετικού υλικού. Πολλές γενετικές ασθένειες με ροπή προς την ανάπτυξη καρκίνου όπως το σύνδρομο Bloom, σχετίζονται με ελαττώματα ή δυσλειτουργίες του συγκεκριμένου επιδιορθωτικού μηχανισμού. Ο ομόλογος ανασυνδυασμός εμφανίζει τέσσερις ξεχωριστούς τρόπους δράσης ανάλογα με την περίπτωση της βλάβης αλλά και το είδος του κυττάρου αν και τα πρώτα στάδια δράσης του είναι παρόμοια σε όλα τα «μονοπάτια». Ο ομόλογος ανασυνδυασμός καταλύεται από μια ομάδα ενζύμων που είναι γνωστά ως «ρεκομπινάσες» (recombinases) [26].

- **DSBR**

Οι λεπτομέρειες του συγκεκριμένου «μονοπατιού» (Double Strand Break Repair) έχουν αναλυθεί λεπτομερώς σε μελέτες του μειωτικού ανασυνδυασμού ο οποίος πραγματοποιείται μετά από προγραμματισμένες δίκλωνες θραύσεις DNA. Το σημείο του χρωμοσώματος στο οποίο υπάρχει η δίκλωνη θραύση επεξεργάζεται για να δώσει μονόκλωνες αλυσίδες DNA (βλέπε σχήμα 3.4 “end resection”). Στην συνέχεια αυτή η αλυσίδα «εισβάλλει» στο ομόλογο υγιές χρωμόσωμα για να αντιγράψει τις γενετικές πληροφορίες που του λείπουν λόγω της θραύσης επομένως σε αυτό το στάδιο έχουμε σύνθεση DNA. Μετά την σύνθεση DNA στο ένα άκρο της αλυσίδας, το δεύτερο άκρο συνδέεται και αυτό με την αλυσίδα του υγιούς χρωμοσώματος και δημιουργεί μια ενδιάμεση δομή με δύο ενώσεις Holliday (Holliday junctions - από το όνομα εκείνου που τις ανακάλυψε πρώτος). Σε αυτό το στάδιο παρατηρείται και πάλι σύνθεση DNA και σύνδεση των άκρων

της νέας αλυσίδας (σχήμα 3.4b). Αφού ολοκληρωθεί η δημιουργία του νέου τμήματος DNA με μοντέλο το γενετικό υλικό του ομόλογου χρωμοσώματος έχουμε την αποσύνδεση των τμημάτων DNA τα οποία συνδέονται στις ενώσεις Holliday. Κατά την αποσύνδεση των αλυσίδων DNA μπορεί να έχουμε εναλλαγή (crossover) τμημάτων των αλληλοεπιδρώντων χρωμοσωμάτων. Τα βασικά ένζυμα-ρεκομπινάσες που συμμετέχουν στον ομόλογο ανασυνδυασμό είναι η RAD50 που δημιουργεί σύμπλοκο με τα MRE11 και NBS1 και συμμετέχουν στην επεξεργασία του DNA για την δημιουργία μονόκλωνης αλυσίδας, οι RAD52 και RAD54 που αλληλεπιδρούν με την RAD50 και τα XRCC2, XRCC3, RAD51B,C,D τα οποία δημιουργούν σύμπλοκα και σταθεροποιούν τους δεσμούς Holliday, ενώ η RAD51C ενδεχομένως να σχετίζεται και με την διάσπαση των δεσμών Holliday [26].

- **SDSA**

Στον μιτωτικό ανασυνδυασμό δεν υπάρχει συσχέτιση με εναλλαγές (crossovers) οπότε έχει προταθεί το μοντέλο SDSA (Synthesis-Dependent Strand-Annealing). Η λειτουργία SDSA (βλέπε σχήμα 3.4c) είναι όμοια με την DSBR στα πρώτα στάδια της επεξεργασίας των άκρων για δημιουργία μονόκλωνων αλυσίδων και στην εισχώρηση στο ομόλογο χρωμόσωμα για την σύνθεση του DNA. Στην συνέχεια όμως αντί να έχουμε σύνδεση και του άλλου άκρου της αλυσίδας στην ενδιάμεση δομή με τις ενώσεις Holliday, έχουμε αποσύνδεση του πρώτου άκρου (αφού ολοκληρωθεί η σύνθεση DNA) και άμεση επανασύνδεση των δύο αυτών άκρων. Το συγκεκριμένο μοντέλο ενδεχομένως να εφαρμόζεται και στις περιπτώσεις του μειωτικού ανασυνδυασμού ο οποίος δεν δίνει εναλλαγές τμημάτων των χρωμοσωμάτων [26].

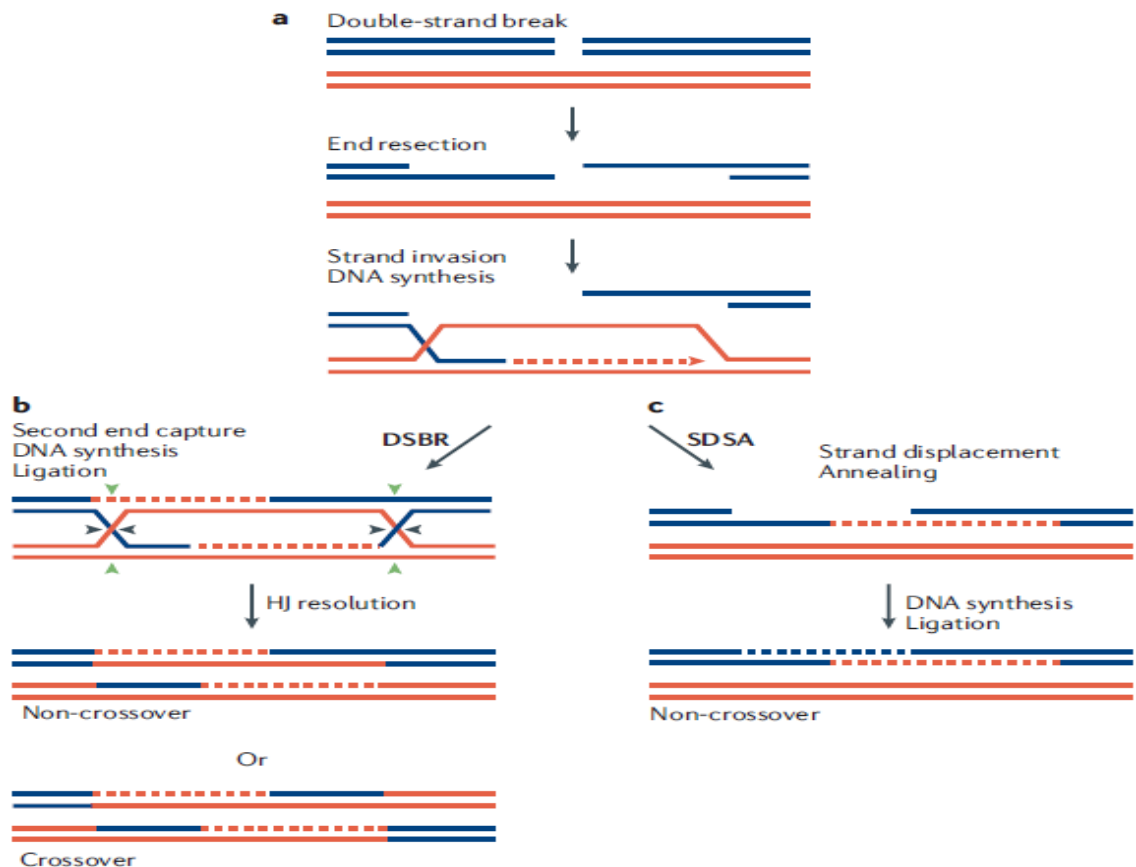
- **SSA**

Αν η δίκλωνη θραύση DNA συμβεί σε περιοχή με επαναλαμβανόμενες αλληλουχίες του γενετικού υλικού τότε μπορεί να επιδιορθωθεί από την διαδικασία ανασυνδυασμού που ονομάζεται SSA (Single-Strand Annealing). Στο συγκεκριμένο «μονοπάτι» τα άκρα στο σημείο της θραύσης επεξεργάζονται για να δώσουν μονόκλωνες αλυσίδες οι οποίες στην συνέχεια επανασυνδέονται μεταξύ

τους. Η SSA δεν απαιτεί όλα τα γονίδια του ομόλογου ανασυνδυασμού αφού η επιδιόρθωση δεν απαιτεί την εισχώρηση σε ομόλογο χρωμόσωμα [26].

- **BIR**

Αν μετά από μια δίκλωνη θραύση υπάρχει μόνο το ένα άκρο της αλυσίδας του DNA (σε περίπτωση που το άλλο μέρος του χρωμοσώματος χαθεί), το μοναδικό άκρο μπορεί να συμμετάσχει σε μια διαδικασία ομόλογου ανασυνδυασμού που ονομάζεται BIR (Break-Induced Replication). Το άκρο επεξεργάζεται όπως και πριν για να δώσει μονόκλωνη αλυσίδα DNA η οποία εισχωρεί και αντιγράφει την γενετική πληροφορία από το ομόλογο χρωμόσωμα. Αν η αλληλουχία που χρησιμοποιείται για την έναρξη του μηχανισμού BIR είναι μέρος μιας οικογένειας επαναλαμβανόμενων αλληλουχιών και βρίσκεται σε μη-ομόλογο χρωμόσωμα, τότε ο BIR μπορεί να έχει σαν αποτέλεσμα μια μη-αμοιβαία μετατόπιση, κατά την οποία τμήμα ενός χρωμοσώματος εισάγεται σε ένα διαφορετικό χρωμόσωμα. Ο BIR εξαρτάται κυρίως από το ένζυμο RAD51 ενώ αποτελεί μέσο για την επιμήκυνση των τελομερών [26].



Σχήμα 3.4: Ο ομόλογος ανασυνδυασμός με τα δυο «μονοπάτια» DSBR και SDSA

2.3.6 Μη-Ομόλογη Ένωση Άκρων (NHEJ)

Η Μη-Ομόλογη Ένωση Άκρων (Non-Homologous End Joining) αποτελεί τον δεύτερο κυτταρικό μηχανισμό για την επιδιόρθωση των δίκλωνων θραύσεων DNA. Η διαφορά του συγκεκριμένου επιδιορθωτικού μηχανισμού με τον προηγούμενο Ομόλογο Ανασυνδυασμό έγκειται στο ότι αυτή την φορά το κύτταρο δεν έχει ομόλογο χρωμόσωμα για να αντιγράψει από αυτό την γενετική πληροφορία που καταστρέφεται κατά την δίκλωνη θραύση. Επί της ουσίας αυτό που καθορίζει το ποιος από τους δύο μηχανισμούς θα χρησιμοποιηθεί για την επιδιόρθωση του DNA είναι η φάση του κυτταρικού κύκλου στην οποία βρίσκεται το κύτταρο όταν εντοπιστεί η ύπαρξη βλάβης. Έτσι αν το κύτταρο βρίσκεται στην S ή στην G2 φάση του κυτταρικού κύκλου και έχει προηγηθεί η αντιγραφή του DNA, με αποτέλεσμα να υπάρχει η λεγόμενη «αδελφή χρωματίδα» (ομόλογο χρωμόσωμα) στον πυρήνα του κυττάρου, θα χρησιμοποιήσει τον Ομόλογο Ανασυνδυασμό για να επιδιορθώσει την δίκλωνη θραύση. Αν όμως το κύτταρο βρίσκεται στην G1 φάση του κυτταρικού κύκλου τότε δεν έχει προλάβει να αντιγράψει το DNA του και θα πρέπει να χρησιμοποιήσει την Μη-Ομόλογη Ένωση Άκρων για την επιδιόρθωση. Η μη ύπαρξη «αδελφής χρωματίδας» για να χρησιμοποιηθεί ως μοντέλο επιδιόρθωσης έχει σαν αποτέλεσμα την πιθανή απώλεια ή και την αλλαγή ορισμένων βάσεων του επιδιορθωμένου DNA και ως εκ τούτου η Μη-Ομόλογη Ένωση Άκρων αν και δρα πολύ γρήγορα είναι επιρρεπής σε λάθη και το επίπεδο πιστότητας της είναι αρκετά μικρό [20]. Ο μηχανισμός NHEJ χωρίζεται σε δύο υποκατηγορίες: τον κλασικό (classical) NHEJ ή cNHEJ και τον εναλλακτικό (alternative) NHEJ ή aNHEJ (επίσης γνωστός ως Backup-NHEJ και Microhomology Mediated End Joining - MMEJ) ο οποίος ανακαλύφθηκε πιο πρόσφατα σε σχέση με τον κλασικό [27].

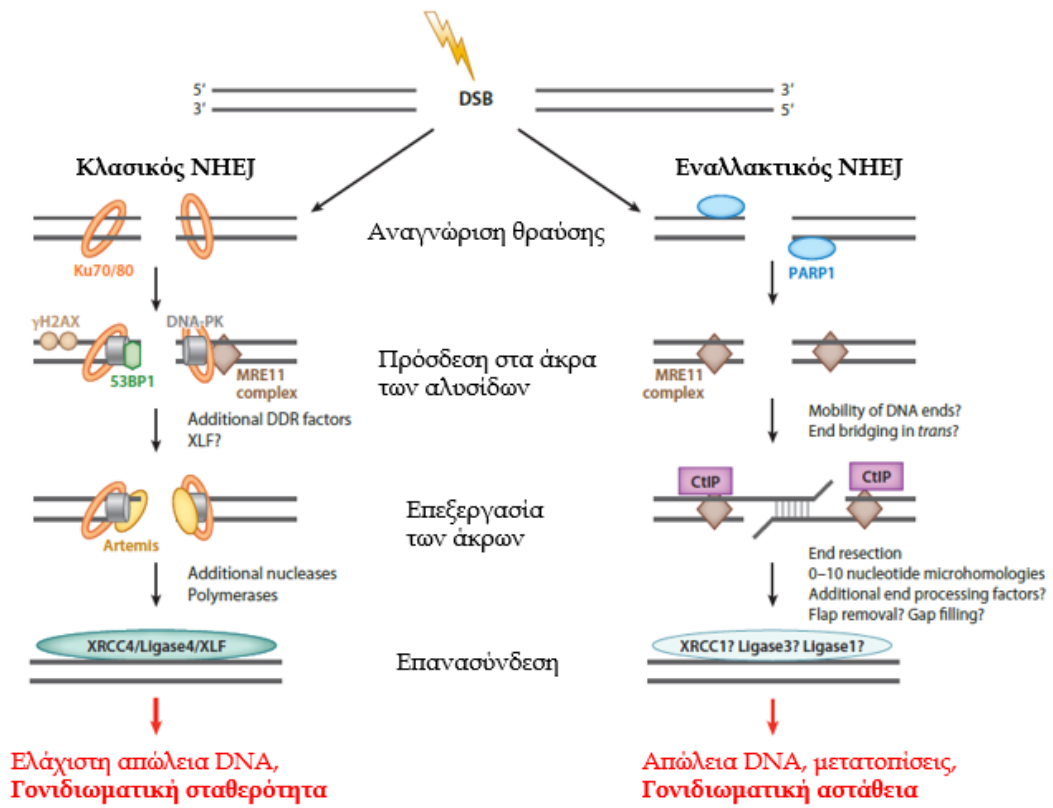
- cNHEJ

Η κλασική Μη-Ομόλογη ένωση άκρων (σχήμα 3.5) φαίνεται σχετικά απλός και άμεσος μηχανισμός: ενώνει τα άκρα της «κομμένης» αλυσίδας DNA χωρίς να απαιτείται η αναζήτηση για το κατάλληλο ομόλογο πρότυπο. Η επιδιόρθωση

ξεκινάει με τον σχηματισμό ενός συμμετρικού ετεροδιμερούς από τις πρωτεΐνες Ku70 και Ku80 οι οποίες κωδικοποιούνται από τα γονίδια XRCC6 και XRCC5 αντίστοιχα. Το ετεροδιμερές που ονομάζεται για συντομία Ku προσδένεται στα σημεία του DNA στα οποία υπάρχει η θραύση και στην συνέχεια έχει την δυνατότητα να μετακινηθεί προς το εσωτερικό της αλυσίδας επιτρέποντας σε περισσότερα ετεροδιμερή Ku να προσδεθούν στο ίδιο μόριο DNA. Μετά την σύνδεση το σύμπλοκο DNA-Ku είναι εξαιρετικά σταθερό και επιτρέπει την πρόσδεση των άλλων επιδιορθωτικών παραγόντων που απαιτούνται όπως η T53BP1, το σύμπλοκο MRE11 και το ένζυμο DNA-PKcs. Το τελευταίο επιτρέπει την φωσφορυλίωση της αλυσίδας του DNA επιτρέποντας έτσι την σύναψη των δύο άκρων. Σε περίπτωση που το σπάσιμο της αλυσίδας είναι πιο σύνθετο (για παράδειγμα όταν έχει δημιουργηθεί από ιοντίζουσα ακτινοβολία) απαιτείται η επιπλέον επεξεργασία των άκρων για να γίνει δυνατή η σύνδεση των δύο τμημάτων. Ένα από τα ένζυμα που πραγματοποιούν αυτή την επεξεργασία είναι το Artemis. Το τελικό στάδιο του μηχανισμού είναι η ένωση των δύο αλυσίδων η οποία πραγματοποιείται από το σύμπλοκο XRCC4/ DNA λιγκάση4/ XLF [27].

- aNHEJ

Η εναλλακτική Μη-Ομόλογη Ένωση Άκρων (σχήμα 3.5) είναι ένας εφεδρικός μηχανισμός για κύτταρα στα οποία υπάρχουν γενετικές ανεπάρκειες για έναν ή περισσότερους από τους παράγοντες που απαιτούνται για τον κλασικό μηχανισμό. Δεν γνωρίζουμε πολλά για τον τρόπο με τον οποίο λειτουργεί ο aNHEJ. Το βασικό χαρακτηριστικό του συγκεκριμένου μηχανισμού είναι ότι είναι επιρρεπής σε λάθη με αποτέλεσμα κατά την επιδιόρθωση να παρατηρούνται εκτεταμένες απώλειες βάσεων και μεταβολές των χρωμοσωμάτων. Έτσι ο aNHEJ είναι πολύ λιγότερο έμπιστος από τον cNHEJ. Πρόσφατες έρευνες δείχνουν ότι η PARP1 και η DNA λιγκάση1 παίζουν σημαντικό ρόλο στην σύνδεση αλυσίδων σε κύτταρα τα οποία εμφανίζουν ανεπάρκεια στο διμερές Ku. Επίσης το σύμπλοκο MRE11 και η CtIP φαίνεται να παίζουν ρόλο στην επεξεργασία των σπασμένων άκρων της αλυσίδας, ενώ η DNA λιγκάση3 και DNA λιγκάση1 μαζί με την XRCC3 ενδεχομένως να πραγματοποιούν την ένωση των αλυσίδων [27].



Σχήμα 3.5: Ο μηχανισμός της Μη-Ομόλογης Ένωσης Άκρων με τις δύο εναλλακτικές εκδοχές του.

Κεφάλαιο 3: Σύνθετες βλάβες DNA

3.1 Η βιολογική σημασία των σύνθετων βλαβών.

Οι σύνθετες - ομαδοποιημένες βλάβες DNA (complex - clustered DNA damage) έχουν μεγάλο βιοφυσικό ενδιαφέρον και αφορούν άμεσα την παρούσα εργασία, αφού τα κύτταρα που εξετάζονται είναι κύτταρα που εμφανίζουν τέτοιες βλάβες στο γενετικό τους υλικό. Ο σκοπός εξάλλου της διπλωματικής είναι η ανίχνευση και επεξεργασία των γονιδίων και των πρωτεϊνών που σχετίζονται με την απόκριση των ανθρώπινων κυττάρων στις σύνθετες βλάβες DNA. Όπως αναφέραμε παραπάνω, βλάβες που προκύπτουν σε μικρή απόσταση μεταξύ τους (μερικά nm) στο μόριο του DNA σχηματίζουν τις ομαδοποιημένες βλάβες ή αλλιώς τις «πολλαπλώς χτυπημένες περιοχές» (multiply damaged sites). Οι πολλαπλώς χτυπημένες περιοχές αποτελούν την «υπογραφή» των ιονιζουσών ακτινοβολιών και είναι γνωστό ότι ο αριθμός των βλαβών και η πολυπλοκότητα τους αυξάνεται όσο αυξάνεται η Γραμμική Εναπόθεση Ενέργειας της ακτινοβολίας [18]. Η μεγάλη βιολογική σημασία αυτών των περιοχών έγκειται στην αδυναμία των κυττάρων να τις επεξεργαστούν και να τις επιδιορθώσουν αποτελεσματικά, σε αντίθεση με τις μεμονωμένες βλάβες που στην πλειονότητα τους επιδιορθώνονται κανονικά από το κύτταρο και το αποτέλεσμα σε περίπτωση λανθασμένης ή και μηδενικής επιδιόρθωσης ποικίλλει από μεταλλάξεις και χρωμοσωμική αστάθεια μέχρι τον κυτταρικό θάνατο [28]. Όταν μια ιοντιζουσα ακτινοβολία διασχίζει ένα κύτταρο εναποθέτει ένα μέρος της ενέργειας που μεταφέρει (ανάλογο της Γραμμικής Εναπόθεσης Ενέργειας της εκάστοτε ακτινοβολίας) στα βιομόρια που αποτελούν το κύτταρο. Ένα από αυτά τα βιομόρια είναι φυσικά και το DNA στο οποίο, λόγω της εναπόθεσης ενέργειας από την ακτινοβολία προκαλούνται διεγέρσεις και ιονισμοί μορίων με αποτέλεσμα την καταστροφή αζωτούχων βάσεων, σπάσιμο των φωσφοδιεστερικών δεσμών στην μια ή και στις δύο αλυσίδες

του DNA κ.α. Αυτές οι βλάβες σχηματίζονται όπως είπαμε σε μια απόσταση μεταξύ μιας ή δυο στροφών της έλικας του DNA, σε αντίθεση με τις βλάβες που σχηματίζονται ενδογενώς (λόγω πχ των μεταβολικών διαδικασιών) και κατανέμονται ομοιόμορφα στο σύνολο του μορίου. Αυτή η χωρική κατανομή των βλαβών, σε συνδυασμό με την δυσκολία του κυττάρου να τις επιδιορθώσει καθιστούν τις ομαδοποιημένες βλάβες τοξικές για το κύτταρο και πολύ συχνά οδηγούν σε προβληματικές καταστάσεις όπως η γήρανση (senescence), η παύση του κυτταρικού κύκλου (cell cycle arrest) και φυσικά ο κυτταρικός θάνατος. Έτσι εξηγείται και η χρήση ιοντίζουσας ακτινοβολίας ως μέσο θεραπείας για καρκινικούς όγκους, αφού βασικός στόχος είναι η θανάτωση των καρκινικών κυττάρων του ασθενούς. Από τις βλάβες που προκαλούνται στο γενετικό υλικό, οι πιο επικίνδυνες για το κύτταρο είναι οι δίκλωνες θραύσεις του DNA που είναι πολύ πιθανό να οδηγήσουν σε κυτταρικό θάνατο λόγω των χαμηλών ποσοστών επιδιόρθωσης. Αντιλαμβάνεται λοιπόν ο αναγνώστης ότι η έρευνα πάνω στις σύνθετες βλάβες του DNA είναι εξαιρετικά σημαντική αφού η κατανόηση των μηχανισμών και η εύρεση πιθανών γονιδίων και πρωτεϊνών που σχετίζονται με αυτές (γονίδια που ενεργοποιούνται ενδεχομένως όταν προκληθούν τέτοιες βλάβες), μπορούν να οδηγήσουν σε σημαντικές ανακαλύψεις τόσο για την ακτινοθεραπεία όσο και για την λειτουργία του κυττάρου γενικότερα [29].

3.2 Μέθοδοι ανίχνευσης σύνθετων βλαβών DNA

Παρόλο που θεωρητικά οι σύνθετες βλάβες στο γενετικό υλικό των κυττάρων είχαν προβλεφθεί από την δεκαετία του '80 η πειραματική ανίχνευση τους δεν ήταν εύκολη και πραγματοποιήθηκε τουλάχιστον μια δεκαετία αργότερα. Οι βασικές μέθοδοι που χρησιμοποιούνται για τον εντοπισμό προβληματικών σημείων στο DNA αξιοποιούν την απόκριση του κυττάρου στην ύπαρξη αυτών των σημείων και κυρίως τους επιδιορθωτικούς μηχανισμούς που ενεργοποιούνται από το κύτταρο. Γνωρίζοντας οι ερευνητές τα ένζυμα και τις πρωτεΐνες που συγκεντρώνονται για την επιδιόρθωση του DNA, έχουν ως στόχο τον εντοπισμό

τους με διάφορους τρόπους, κάτι που θα υποδείξει την ύπαρξη βλαβών. Η πρώτη μέθοδος που χρησιμοποιήθηκε για την ανίχνευση σύνθετων βλαβών στο DNA ήταν αυτή της ηλεκτροφόρησης. Κατά την ηλεκτροφόρηση διοχετεύεται ηλεκτρικό ρεύμα μέσω ηλεκτροδίων σε ένα μέσο (συνήθως τζελ) που πάνω του έχει τοποθετηθεί σε ένα σημείο το προς ανάλυση δείγμα. Το αποτέλεσμα είναι ότι τα φορτισμένα σωματίδια κινούνται προς τα ηλεκτρόδια με ταχύτητες διαφορετικές ανάλογα με το φορτίο τους και αντιστρόφως ανάλογες με το μέγεθος τους. Έτσι τα περισσότερα φορτισμένα και μικρότερα μόρια απομακρύνονται περισσότερο από το αρχικό σημείο, ενώ τα μεγαλύτερα και λιγότερο φορτισμένα λιγότερο, οπότε επέρχεται διαχωρισμός. Στην περίπτωση μας χρησιμοποιούνται επιδιορθωτικά ένζυμα (όπως τα ανθρώπινα OGG1 και APE1) ως «δείκτες» των βλαβών, αφού όπου υπάρχει ζημιά στο μόριο του DNA συγκεντρώνονται τα συγκεκριμένα ένζυμα για να την επιδιορθώσουν και έτσι με την τεχνική της ηλεκτροφόρησης πετυχαίνουμε την ποσοτικοποίηση αυτών των βλαβών [30]. Μια άλλη προσέγγιση για την ανίχνευση αβασικών σημείων στο μόριο του DNA λόγω έκθεσης σε ιοντίζουσα ακτινοβολία είναι αυτή του ανοσοσήμανσης, με διάφορες τεχνικές όπως του ανοσοφθορισμού και της ανοσοστόπωσης κατά Western (Western Blotting) και όλες χρησιμοποιούν κατάλληλα αντισώματα ώστε να ανιχνευθεί μία συγκεκριμένη πρωτεΐνη που σχετίζεται με συγκεκριμένη βλάβη στο DNA. Ένα μεγάλο μειονέκτημα των ηλεκτροφορητικών μεθόδων είναι ότι δεν ανιχνεύουν μικρά τμήματα γενετικού υλικού (μικρότερα των 5 χιλιάδων ζευγών βάσεων), τα οποία προκύπτουν συχνά κατά την ακτινοβόληση με ακτινοβολία υψηλής LET (σύνθετες δίκλωνες θραύσεις) με αποτέλεσμα να υποτιμάται ο πραγματικός αριθμός των δίκλωνων θραύσεων του DNA. Μια πιο ευαίσθητη μέθοδος ανίχνευσης δίκλωνων θραύσεων στο γενετικό υλικό είναι αυτή της ιστόνης H2AX [31]. Η H2AX ανήκει στην οικογένεια πρωτεϊνών H2A και αποτελεί μέρος του οκταμερούς των ιστονών που συνθέτουν τα νουκλεοσώματα. Όταν δημιουργούνται δίκλωνες θραύσεις στο DNA ακολουθεί πάντα η φωσφορυλίωση της συγκεκριμένης ιστόνης και αποτελεί σήμα για το κύτταρο ώστε να ενεργοποιήσει τους επιδιορθωτικούς μηχανισμούς. Η φωσφορυλιωμένη πρωτεΐνη ονομάζεται γ -H2AX και όπου υπάρχει δίκλωνη θραύση στο γενετικό υλικό

δημιουργούνται μικρές εστίες της. Επιπλέον αυτές οι εστίες της γ -H2AX αντιστοιχούν σε δίκλωνες θραύσεις σε αναλογία ένα προς ένα και χρησιμοποιούνται σαν βιοδείκτης για τις βλάβες με τις κατάλληλες μεθόδους οπτικοποίησης [32].

3.3 Σύνθετες βλάβες DNA και συστημικά φαινόμενα

Εκτός από τις άμεσες επιδράσεις των σύνθετων βλαβών του γενετικού υλικού στο κύτταρο, προκαλούνται στον οργανισμό φαινόμενα που σχετίζονται με την απόκριση του ανοσοποιητικού συστήματος στην διατάραξη της ομαλής λειτουργίας του οργανισμού και τα οποία ονομάζουμε συστημικά φαινόμενα. Είναι γνωστό ότι όταν προκαλείται σύνθετη ζημιά στο γενετικό υλικό που συνήθως ακολουθείται από γονιδιωματική αστάθεια, ενεργοποιούνται οι μηχανισμοί επιδιόρθωσης του DNA (DNA damage response - DDR). Οι μηχανισμοί αυτοί μπορούν επίσης να ενεργοποιήσουν το ανοσοποιητικό σύστημα του οργανισμού και το σήμα για την ενεργοποίηση δίνεται από διαφορετικά στοιχεία του DDR όπως οι αισθητήρες βλάβης DNA, ορισμένες επιδιορθωτικές πρωτεΐνες και κάποιες κινάσες (ένζυμα που καταλύουν την μεταφορά φωσφορικών ομάδων) [33]. Για παράδειγμα έχει αποδειχθεί ότι ανθρώπινα κύτταρα (HCA2 ινοβλάστες) που έχουν ακτινοβοληθεί με ακτίνες - X παρουσιάζουν φλεγμονώδη αντίδραση με την έκκριση κυτοκινών (μικρές πρωτεΐνες που παίζουν σημαντικό ρόλο στην «επικοινωνία» των κυττάρων μέσω της μεταφοράς τους από το ένα κύτταρο στο άλλο) όπως η ιντερλευκίνη-6 [34]. Τα συστημικά φαινόμενα αποτελούν μέρος της άμυνας του οργανισμού που επιχειρεί να απομακρύνει τα χτυπημένα και δυσλειτουργικά κύτταρα και να διατηρήσει την ομαλή λειτουργία των ιστών. Σε άλλη περίπτωση έχει δείχθει ότι για ασθενείς με το σύνδρομο ataxia-telangiectasia στο οποίο η επιδιορθωτική πρωτεΐνη ATM είναι ελαττωματική, μικρά θραύσματα DNA συσσωρεύονται στο κυτταρόπλασμα των κυττάρων του ασθενούς. Τα θραύσματα αυτά εντοπίζονται από τις αναγνωριστικές μονάδες του

ανοσοποιητικού συστήματος, που η κανονική τους λειτουργία είναι να εντοπίζουν DNA από ιούς που εισβάλλουν στο κύτταρο. Το αποτέλεσμα είναι να σημαίνει λάθος «συναγερμός» ικής εισβολής και να παράγεται η ιντερφερόνη τύπου I [35]. Στην κατηγορία των συστημικών φαινομένων ανήκουν και οι λεγόμενες «μη-στοχευμένες επιδράσεις της ακτινοβολίας» (non - targeted effects / NTE). Κατά την ακτινοβολήση ενός οργανισμού με ιοντίζουσα ακτινοβολία είναι πιθανό να προκύψουν επιδράσεις σε περιοχές οι οποίες δεν έχουν δεχθεί ακτινοβολία και οι οποίες οφείλονται στην ανοσολογική απόκριση του οργανισμού. Οι μη-στοχευμένες επιδράσεις της ακτινοβολίας μπορούν να χωριστούν σε δύο κατηγορίες, τις «κοντινές» ή παρακείμενες (bystander) που παρατηρούμε μη - ακτινοβολημένα κύτταρα να παρουσιάζουν συμπεριφορά παρόμοια με τα γειτονικά τους που έχουν δεχθεί την ακτινοβολία και τις «μακρινές» όπως είναι η «απόκριση απομακρυσμένου ιστού σε ακτινοβολία» (abscopal effect). Το τελευταίο είναι μια σπάνια κατάσταση κατά την θεραπεία μεταστατικών καρκίνων όπου παρατηρείται σμίκρυνση όγκων που δεν δέχονται θεραπεία ταυτόχρονα με την σμίκρυνση εκείνων που δέχονται την εντοπισμένη θεραπεία. Οι μη - στοχευμένες επιδράσεις, οφείλονται όπως είπαμε στην αντίδραση του ανοσοποιητικού συστήματος του οργανισμού και συνήθως περιλαμβάνουν την απελευθέρωση διάφορων χημικών και βιολογικών παραγόντων από τα ακτινοβολημένα κύτταρα με αποτέλεσμα την «επικοινωνία» της «επίθεσης» που δέχονται και στα υπόλοιπα κύτταρα του οργανισμού [28].

Κεφάλαιο 4 Στοιχεία Βιοπληροφορικής

4.1 Εισαγωγή - Η σημασία της βιοπληροφορικής

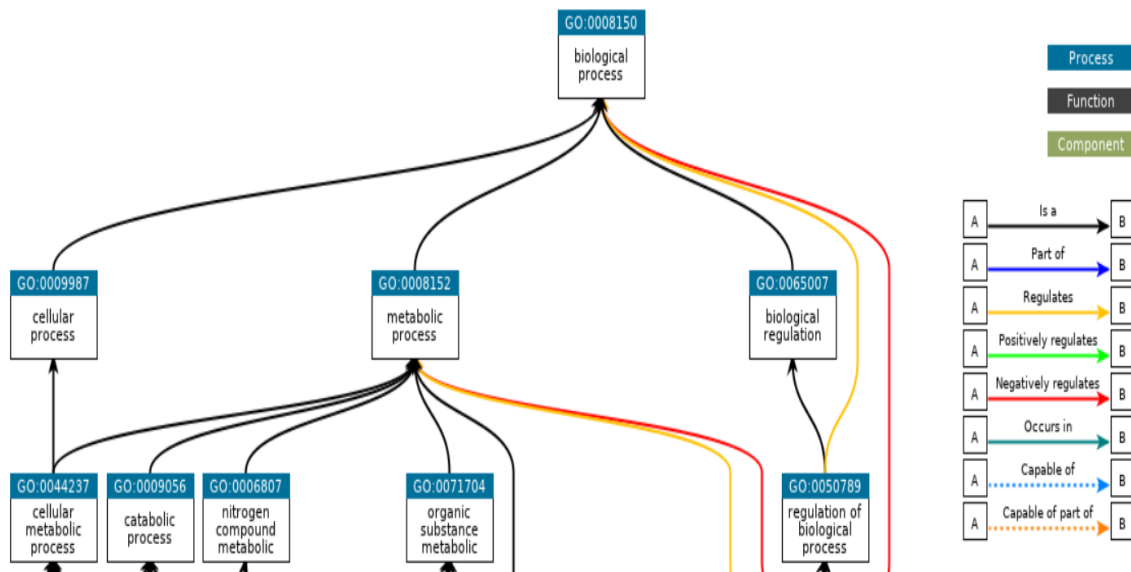
Η βιοπληροφορική είναι ένας διεπιστημονικός τομέας που συνδέει την επιστήμη των υπολογιστών με την βιολογία. Ορίζεται ως η αξιοποίηση της τεχνολογίας των ηλεκτρονικών υπολογιστών για την αποθήκευση, ανάκτηση, επεξεργασία και διανομή πληροφοριών που σχετίζονται με τα βιολογικά μακρομόρια όπως το DNA, το RNA και οι πρωτεΐνες [36]. Σχεδόν όλες οι βιολογικές έρευνες σήμερα σε κάποιο σημείο της πορείας για την απάντηση ενός ερωτήματος, απαιτούν την εφαρμογή κάποιου μαθηματικού, στατιστικού ή υπολογιστικού εργαλείου ως επιπλέον βοήθεια για την επεξεργασία και την «σύνθεση» των καταγεγραμμένων δεδομένων και πληροφοριών. Ανεξάρτητα από το αν η εφαρμογή αυτή είναι απλή ή σύνθετη είναι πλέον ξεκάθαρο ότι μαθηματικά και υπολογιστικά εργαλεία έχουν γίνει αναπόσπαστο κομμάτι της σύγχρονης βιολογικής έρευνας [37]. Για την συγκεκριμένη διπλωματική που ο σκοπός είναι η ανίχνευση των γονιδίων που σχετίζονται με την απόκριση στις σύνθετες βλάβες DNA με χρήση εργαλείων βιοπληροφορικής ανάλυσης, εύκολα αντιλαμβάνεται ο αναγνώστης ότι το πιο σημαντικό ρόλο στην εκπόνηση της εργασίας παίζει η επιστήμη της βιοπληροφορικής. Η ανάπτυξη της βιοπληροφορικής σαν επιστήμη έρχεται σαν αποτέλεσμα της προόδου που έχει επέλθει τα τελευταία 30 - 40 χρόνια τόσο στην μοριακή βιολογία όσο και στην επιστήμη των υπολογιστών. Ο βασικός λόγος για τον οποίο έγινε τόσο «διάσημη» ως επιστήμη, ήταν η πρόοδος στην μελέτη του γονιδιώματος που παρήγαγε μια άνευ προηγουμένου ποσότητα βιολογικών δεδομένων. Αυτή η «έκρηξη» πληροφορίας δημιούργησε την άμεση ανάγκη για αποδοτικά εργαλεία επεξεργασίας και ανάλυσης των δεδομένων. Η έμφαση δίνεται στην χρήση υπολογιστών επειδή η ανάλυση των γενετικών δεδομένων και η αλληλουχία του γονιδιώματος ενός οργανισμού εμφανίζει μεγάλη

επαναληψιμότητα και μαθηματική πολυπλοκότητα. Χαρακτηριστικό παράδειγμα χρήσης της επιστήμης των υπολογιστών στην βιολογική έρευνα αποτελεί η μοντελοποίηση συστημάτων νουκλεϊκών οξέων [38]. Ο απώτερος στόχος της βιοπληροφορικής είναι η καλύτερη κατανόηση του κυττάρου και των λειτουργιών του σε μοριακό επίπεδο. Αναλύοντας την μοριακή αλληλουχία και δομή, η βιοπληροφορική παράγει νέα γνώση και μια πιο πλήρη εικόνα του κυττάρου [37].

4.2 Όροι γονιδιακής οντολογίας

Η Γονιδιακή Οντολογία (Gene Ontology - GO) είναι ένα μεγάλο πρότζεκτ της βιοπληροφορικής που έχει σαν στόχο να διευκολύνει και να ενοποιήσει την αναπαράσταση γονιδίων και πρωτεϊνών ανάλογα με τις λειτουργίες που αυτά επιτελούν [39]. Παρέχει στην επιστημονική κοινότητα ένα δομημένο «λεξιλόγιο» που χρησιμοποιείται για την επισήμανση γονιδίων και γονιδιακών προϊόντων (δηλαδή πρωτεΐνες και RNA) με σκοπό την εύχρηστη περιγραφή των χαρακτηριστικών λειτουργιών των παραπάνω, κατηγοριοποιημένων σε τρεις μεγάλες «βιολογικές ενότητες» που είναι κοινές για όλους τους οργανισμούς: «μοριακή λειτουργία» (molecular function), «βιολογική διαδικασία» (biological process) και «κυτταρικό συστατικό» (cellular component) . Οι όροι οντολογίας έχουν επεκταθεί ώστε να περιέχουν σχεδόν όλες τις περιοχές της βιολογίας και βιολογικά «μονοπάτια» ενώ έχει γίνει και μεγάλη προσπάθεια να υπάρξει αναλυτική σήμανση των ορθόλογων γονιδίων (δηλαδή γονίδια που συναντώνται σε διαφορετικούς οργανισμούς αλλά προέρχονται από το γονιδίωμα του κοινού προγόνου αυτών των δύο οργανισμών [40]) μεταξύ ενός μεγάλου αριθμού γονιδιωμάτων αναφοράς μεταξύ των οποίων του ανθρώπου και άλλων σημαντικών οργανισμών. Κάθε όρος γονιδιακής οντολογίας δίνεται με ένα αλφαριθμητικό το οποίο αποτελεί την «ταυτότητα» του συγκεκριμένου όρου, ένα όνομα και μια από τις τρεις ενότητες που αναφέραμε, στην οποία ανήκει ο όρος. Για παράδειγμα ένας όρος γονιδιακής οντολογίας που περιγράφει μια λειτουργία της πρωτεΐνης λακτάσης είναι ο εξής: id: GO:0000016 (η ταυτότητα του όρου),

name: lactase activity, ontology: molecular_function. Η δομή της GO μπορεί να αναπαρασταθεί με ένα γράφημα (εικόνα 4.1), όπου κάθε όρος γονιδιακής οντολογίας είναι ένας κόμβος και οι σχέσεις μεταξύ των όρων αναπαρίστανται με βέλη. Έτσι δημιουργείται το λεγόμενο δέντρο των προγόνων, όπου για κάθε όρο μπορείς να βρεις τους πιο εξειδικευμένους όρους που αποτελούν υποσύνολο του αρχικού (child terms) ή και τον πιο γενικό όρο ο οποίος περιλαμβάνει τον αρχικό (parent terms).



Εικόνα 4.1: Ένα τυπικό δέντρο προγόνων για όρους γονιδιακής οντολογίας

Η GO είναι ένα δυναμικό πρότζεκτ το οποίο συνεχώς μεταβάλλεται καθώς προστίθενται νέοι όροι ενώ τροποποιούνται και διορθώνονται οι ήδη υπάρχοντες καθώς η βιολογική έρευνα προχωράει και εξελίσσεται. Η πρόσβαση στα δεδομένα της GO είναι δωρεάν και ανοιχτή μέσω του επίσημου σαιτ <http://geneontology.org/> αλλά και του <http://amigo.geneontology.org/> το οποίο περιέχει εργαλεία για την επεξεργασία των δεδομένων [39].

4.3 Βάσεις δεδομένων

Βασικό ρόλο σε μια βιοπληροφορική ανάλυση παίζουν οι βάσεις δεδομένων από όπου μπορεί κανείς να συγκεντρώσει δεδομένα τα οποία στην συνέχεια επεξεργάζεται για να καταλήξει σε κάποια συμπεράσματα. Οι ερευνητές που θα εκτελέσουν ένα πείραμα, για παράδειγμα μια ακτινοβολήση ενός συγκεκριμένου τύπου κυττάρων, θα συγκεντρώσουν τα αποτελέσματα του πειράματος και θα τα καταθέσουν σε κάποια βάση δεδομένων ώστε να είναι διαθέσιμα για επεξεργασία από οποιονδήποτε ενδιαφερόμενο. Στην συγκεκριμένη εργασία χρησιμοποιήσαμε ορισμένες από τις πιο γνωστές βάσεις δεδομένων τις οποίες θα παρουσιάσουμε εδώ:

- Η NCBI (National Center for Biotechnology Information) [41] είναι μια από τις μεγαλύτερες και πιο σημαντικές βάσεις δεδομένων. Δημιουργήθηκε το 1988 με σκοπό την ανάπτυξη νέων προγραμματιστικών τεχνολογιών για την καλύτερη κατανόηση των μοριακών και γενετικών διαδικασιών που καθορίζουν τις ασθένειες αλλά και τους υγιείς οργανισμούς. Η NCBI είναι υπεύθυνη για την δημιουργία αυτοματοποιημένων συστημάτων για την αποθήκευση και ανάλυση γνώσεων σχετικών με την μοριακή βιολογία, την βιοχημεία και την γενετική. Επιπλέον έχει σαν στόχο την διευκόλυνση της ερευνητικής και ιατρικής κοινότητας ως προς την χρήση των δεδομένων της καθώς και την καθοδήγηση των προσπαθειών για συγκέντρωση βιοτεχνολογικών πληροφοριών σε εθνικό (ΗΠΑ) και διεθνές επίπεδο. Στην NCBI περιλαμβάνονται πολλές μικρότερες βάσεις δεδομένων όπως η Gene και η PubMed οι οποίες χρησιμοποιήθηκαν στην παρούσα εργασία.
- Η AmiGO2 [42] που αναφέρθηκε και στην προηγούμενη παράγραφο είναι μια διαδικτυακή εφαρμογή που δημιουργήθηκε από την ομάδα του GeneOntology project με σκοπό να διευκολύνει την άντληση δεδομένων από ενδιαφερόμενους ερευνητές. Επιτρέπει στον χρήστη την εύκολη αναζήτηση, την ταξινόμηση, την οπτικοποίηση και το κατέβασμα (download) των δεδομένων που τον ενδιαφέρουν σχετικά με όρους οντολογίας κατά βάση αλλά και γονίδια και προϊόντα γονιδίων

- Η Ensembl genome database program [43] είναι ένα κοινό επιστημονικό πρόγραμμα ανάμεσα στο ευρωπαϊκό ινστιτούτο βιοπληροφορικής και του Wellcome Trust Sanger institute με σκοπό αντίστοιχο με αυτόν της NCBI, δηλαδή την παροχή δεδομένων γενετικού και βιολογικού περιεχομένου. Η Ensembl παρέχει μια κεντρική πηγή άντλησης δεδομένων για γενετιστές, μοριακούς βιολόγους και άλλους ερευνητές που ερευνούν το γονιδίωμα του ανθρώπου και άλλων οργανισμών. Επιπλέον στην ιστοσελίδα της Ensembl υπάρχει ειδική πλατφόρμα για την μετατροπή της ονοματολογίας των γονιδίων πχ από ensembl ID στην επίσημη ονομασία των γονιδίων (official gene symbol) την οποία και χρησιμοποιήσαμε στην εργασία αυτή.

4.4 Πλατφόρμες Βιοπληροφορικής Ανάλυσης

Σε αυτή την ενότητα θα παρουσιάσουμε τις πλατφόρμες που χρησιμοποιήθηκαν στην παρούσα εργασία, με σκοπό την ανάλυση των πληροφοριών που αντλήσαμε από τις βάσεις δεδομένων που ήδη αναφέραμε. Πρόκειται για εργαλεία (tools) βιοπληροφορικής ανάλυσης τα οποία είναι διαθέσιμα στο διαδίκτυο και που μπορεί ο καθένας να χρησιμοποιήσει δωρεάν για να πραγματοποιήσει την μελέτη που τον ενδιαφέρει.

4.4.1 Database for Annotation, Visualization and Integrated Discovery (DAVID)

Μια από τις βασικές πλατφόρμες που χρησιμοποιήθηκαν είναι η DAVID [44]. Η συγκεκριμένη πλατφόρμα χρησιμοποιείται κυρίως για το λεγόμενο pathway enrichment analysis, δηλαδή την εύρεση των βιολογικών «μονοπατιών» στα οποία

ανήκουν τα γονίδια που μας ενδιαφέρουν. Για να ξεκινήσει κάποιος να δουλεύει στην DAVID πρέπει αρχικά να ανεβάσει (upload) μια λίστα με ονόματα γονιδίων τα οποία στην συνέχεια μπορεί να επεξεργαστεί. Αφού ο χρήστης δώσει τα ονόματα των γονιδίων το σύστημα ελέγχει ότι είναι έγκυρα και στην συνέχεια τον προτρέπει να επιλέξει τον επιθυμητό οργανισμό στον οποίο ανήκουν τα δοθέντα γονίδια. Στην συνέχεια ο χρήστης έχει να επιλέξει ανάμεσα σε μια σειρά εργαλείων. Μπορεί να ζητήσει αλλαγή της βάσης ονοματολογίας των γονιδίων πηγαίνοντας για παράδειγμα από την επίσημη ονοματολογία των γονιδίων σε ονοματολογία κατά ensembl ή κατά HUGO. Αφού βρει την επιθυμητή βάση ονοματολογίας επιλέγει ανάμεσα σε κατηγορίες σήμανσης των γονιδίων όπως pathways ή gene ontology που δίνει τις βιολογικές λειτουργίες που σχετίζονται με τα γονίδια, diseases που επιστρέφει τις ασθένειες που σχετίζονται με τα γονίδια κ.α. Κάθε ενότητα αποτελείται από επιμέρους βάσεις δεδομένων από τις οποίες τα εργαλεία του DAVID αντλούν τα αποτελέσματα. Για παράδειγμα η κατηγορία pathways έχει αρκετές βάσεις δεδομένων σαν επιλογές όπως την kegg pathways, την biocarta κ.α. ενώ η κατηγορία disease έχει τις gad disease, omim disease κ.α. Στην παρούσα εργασία με την βοήθεια του DAVID κάναμε pathway enrichment analysis στα γονίδια που μας ενδιέφεραν επιλέγοντας την ενότητα gene ontology και την βάση δεδομένων “goterm bp all”.

4.4.2 STRING

Το εργαλείο STRING βρίσκει τα protein – protein interactions από ένα σύνολο γονιδίων [45]. Αυτό σημαίνει ότι του δίνεις ένα σύνολο γονιδίων σαν input και το output είναι μια εικόνα – χάρτης που δίνει τα γονίδια που είναι γνωστό ότι αλληλεπιδρούν μεταξύ τους, είτε αυτό έχει επιβεβαιωθεί πειραματικά είτε μέσω άλλων παραμέτρων όπως το να βρίσκονται σε γειτονικές περιοχές στο χρωμόσωμα κ.α. Η πλατφόρμα STRING λειτουργεί ως εξής: ο χρήστης αρχικά εισάγει μια λίστα με τα ονόματα των γονιδίων που τον ενδιαφέρουν και επιλέγει τον οργανισμό που τον ενδιαφέρει και στον οποίο ανήκουν τα γονίδια. Στην συνέχεια το σαι επιστρέφει μια διαδραστική εικόνα με τα γονίδια να απεικονίζονται σαν μικρές μπάλες και την οποία ο χρήστης μπορεί να επεξεργαστεί επιλέγοντας μια σειρά

από παραμέτρους όπως το σκορ αλληλεπίδρασης (πόσο σίγουρη είναι η συγκεκριμένη αλληλεπίδραση) και την πηγή που επιβεβαιώνει την συγκεκριμένη αλληλεπίδραση για παράδειγμα πειραματικά επιβεβαιωμένη αλληλεπίδραση, γειτνίαση γονιδίων, συν - έκφραση κ.α.

4.4.3 DisGeNET

Η διαδικτυακή πλατφόρμα DisGeNET περιλαμβάνει μια από τις μεγαλύτερες συλλογές γονιδίων που σχετίζονται με ανθρώπινες ασθένειες [46]. Συγκεντρώνει δεδομένα από άλλες μικρότερες βάσεις δεδομένων καθώς και από την επιστημονική βιβλιογραφία και έτσι πετυχαίνει να είναι μια από τις πληρέστερες μηχανές αναζήτησης για ασθένειες και σύνδρομα που σχετίζονται με μεταλλάξεις ή γενικότερα βλάβες σε κάποιο/α γονίδιο/α. Ο τρόπος λειτουργίας της συγκεκριμένης πλατφόρμας είναι αρκετά απλός και φιλικός για τον χρήστη. Στην αρχική σελίδα ο χρήστης δίνει το όνομα του γονιδίου που τον ενδιαφέρει ή αν αυτά είναι πολλά έχει την δυνατότητα να τα εισάγει όλα μαζί υπό μορφή λίστας. Στην συνέχεια πρέπει να επιλέξει αν θέλει μια γενική σύνοψη των αποτελεσμάτων, στην οποία εμφανίζονται όλες οι ασθένειες που σχετίζονται με το συγκεκριμένο γονίδιο μαζί με έναν δείκτη για την κάθε συσχέτιση. Ο δείκτης αυτός παίρνει τιμές από μηδέν έως ένα και όσο πιο κοντά στην μονάδα είναι ο δείκτης τόσο πιο σίγουρη είναι η συγκεκριμένη συσχέτιση. Επιπλέον στην σύνοψη των αποτελεσμάτων φαίνεται ο αριθμός των δημοσιεύσεων που υποστηρίζουν την εκάστοτε συσχέτιση καθώς και οι χρονολογίες που αναφέρθηκαν πρώτη και τελευταία φορά η συσχέτιση του γονιδίου με την ασθένεια. Η άλλη επιλογή του χρήστη πέρα από την σύνοψη είναι να δει τα στοιχεία που υποστηρίζουν την συσχέτιση, όπου στην ιστοσελίδα που ανοίγει εμφανίζονται, πέρα από την ασθένεια και τον δείκτη που αναφέραμε παραπάνω, ο τύπος της συσχέτισης (πχ βιοδείκτης, αλλαγή έκφρασης του γονιδίου κ.α.), η βάση δεδομένων που υποστηρίζει την συσχέτιση καθώς και το ID της δημοσίευσης αλλά και τη συγκεκριμένη πρόταση μέσα στην δημοσίευση που αυτή αναφέρεται. Στην σύνοψη των αποτελεσμάτων ο χρήστης έχει την δυνατότητα να εισάγει φίλτρα,

όπως την ελάχιστη επιτρεπτή τιμή που επιθυμεί να παίρνει ο δείκτης για μια δεδομένη συσχέτιση (κρατώντας έτσι μόνο εκείνες τις ασθένειες για τις οποίες είναι πολύ πιθανή η συσχέτιση τους με το εκάστοτε γονίδιο), τον ελάχιστο αριθμό των δημοσιεύσεων που αναφέρουν την συσχέτιση κ.α. Επίσης στην πλατφόρμα αυτή μπορεί να γίνει και η αντίστροφη διαδικασία, δηλαδή ο χρήστης να εισάγει μια συγκεκριμένη ασθένεια ή κάποιο σύνδρομο και η πλατφόρμα του επιστρέφει όλα τα γονίδια που σχετίζονται με αυτήν.

4.4.4 Bioinformatics and Evolutionary Genomics (Van de Peer Lab)

Η συγκεκριμένη ιστοσελίδα ανήκει σε μια ερευνητική ομάδα βιοπληροφορικής με έδρα την Γάνδη του Βελγίου. Μέσα από το σαιτ αυτό μπορεί κανείς να αποκτήσει πρόσβαση σε διάφορα λογισμικά που έχουν αναπτυχθεί με την βοήθεια της συγκεκριμένης ομάδας. Ένα από αυτά είναι και το VENN DIAGRAMS (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) το οποίο χρησιμοποιήσαμε στην παρούσα εργασία. Όπως υποδεικνύει και το όνομα του λογισμικού, η λειτουργία του είναι η κατασκευή διαγραμμάτων Venn τα οποία είναι πολύ χρήσιμα στην επιστήμη της βιοπληροφορικής για την σύγκριση μεγάλων συνόλων δεδομένων. Η λειτουργία του είναι εξαιρετικά απλή αφού ο χρήστης απλά ανεβάζει τα αρχεία που επιθυμεί να συγκρίνει (πχ δυο λίστες γονιδίων) και το λογισμικό επιστρέφει το διάγραμμα Venn που δείχνει την επικάλυψη που υπάρχει μεταξύ των αρχείων, δηλαδή πόσα και ποια στοιχεία ανήκουν αποκλειστικά σε ένα αρχείο και πόσα ανήκουν σε παραπάνω από ένα αρχείο.

4.4.5 Gene Expression Profiling Interactive Analysis (GEPIA)

Το GEPIA είναι μια διαδικτυακή πλατφόρμα για την ανάλυση της έκφρασης γονιδίων σε διάφορες μορφές καρκίνου σε σύγκριση με την έκφραση τους σε φυσιολογικούς ιστούς [47]. Δηλαδή για κάθε γονίδιο που μας ενδιαφέρει,

επιλέγοντας διαφορετικούς τύπους καρκίνου, το σαιτ επιστρέφει ένα διάγραμμα από το οποίο μπορούμε να αποφανθούμε αν η αυξημένη έκφραση του (έκφραση σε τιμές μεγαλύτερες του φυσιολογικού) που οφείλεται στην ύπαρξη καρκινικού ιστού, έχει στατιστικά σημαντικό ρόλο στην επιβίωση του ασθενούς. Ο χρήστης αρχικά δίνει το όνομα ενός γονιδίου και στην συνέχεια έχει από το σαιτ ορισμένες επιλογές ανάλογα με την ανάλυση που επιθυμεί να κάνει. Επιλέγοντας τα λεγόμενα survival plots μεταφέρεται σε μια νέα ιστοσελίδα στην οποία θέτει ορισμένες παραμέτρους και επιλέγει τον τύπο καρκίνου που τον ενδιαφέρει. Το αποτέλεσμα είναι ένα διάγραμμα στο οποίο απεικονίζεται το ποσοστό επιβίωσης ασθενών, βασισμένο σε πραγματικά δεδομένα στην κλίμακα του χρόνου ανάλογα με το αν η έκφραση του συγκεκριμένου γονιδίου ήταν υψηλή ή χαμηλή καθώς και κάποιοι δείκτες (p-value) που δίνουν την στατιστική σημασία των αποτελεσμάτων. Δηλαδή μέσω αυτών των δεικτών ο χρήστης καταλαβαίνει αν η διαφορετική έκφραση του κάθε γονιδίου παίζει πραγματικά ρόλο στην επιβίωση των ασθενών, κάτι που ενδεχομένως θα μπορούσε να χρησιμοποιηθεί ακόμα και σαν διαγνωστικός παράγοντας ή αν τα συγκεκριμένα αποτελέσματα θα μπορούσαν να είναι και τυχαία και να μην υπάρχει πραγματική συσχέτιση μεταξύ του γονιδίου και του τύπου καρκίνου που τον ενδιαφέρει.

4.4.6 SeneQuest

Το συγκεκριμένο εργαλείο αποτελεί μια προσπάθεια να συγκεντρωθεί σε μια ενιαία βάση δεδομένων οι μέχρι τώρα γνώσεις μας για την συσχέτιση γονιδίων και κυτταρικής γήρανσης (cellular senescence) [48]. Η κυτταρική γήρανση είναι μια κατάσταση στην οποία μπορεί να περιέλθει ένα κύτταρο μετά από κάποιο στρεσογόνο γεγονός ή μέσω ορισμένων φυσιολογικών διαδικασιών και χαρακτηρίζεται από μη - αναστρέψιμη διακοπή του κυτταρικού κύκλου με ταυτόχρονη έκκριση μορίων (που ενδεχομένως λειτουργούν ως σήμα για τα περιβάλλοντα κύτταρα), καταστροφή μακρομορίων στο εσωτερικό του και αλλαγή των μεταβολικών διαδικασιών [48]. Όπως έχουμε ήδη αναφέρει η ακτινοβολία κυττάρων με ιοντίζουσα ακτινοβολία και η επακόλουθη καταστροφή του

γενετικού τους υλικού είναι ένας από τους παράγοντες που μπορεί να οδηγήσουν το κύτταρο σε αυτή την κατάσταση. Στην πλατφόρμα αυτή που έφτιαξε η ερευνητική ομάδα του καθηγητή κύριου Γοργούλη, συγκεντρώνεται από την διεθνή βιβλιογραφία η πληροφορία για τη σύνδεση ανάμεσα στην αλλαγή έκφρασης διαφόρων γονιδίων και την κατάσταση της κυτταρικής γήρανσης. Δηλαδή ποια γονίδια αλλάζουν τον ρυθμό έκφρασης τους (είτε υψηλότερος είτε χαμηλότερος) όταν το κύτταρο βρίσκεται ή πρόκειται να βρεθεί σε αυτή την κατάσταση. Η λειτουργία του είναι εξαιρετικά απλή: ο χρήστης απλά δίνει το όνομα του γονιδίου και η πλατφόρμα επιστρέφει σαν αποτέλεσμα αν και πόσες φορές έχει αναφερθεί σε κάποια δημοσίευση η συσχέτιση της αυξημένης ή μειωμένης έκφρασης του γονιδίου αυτού με την κυτταρική γήρανση.

Β' Υπολογιστικό Μέρος

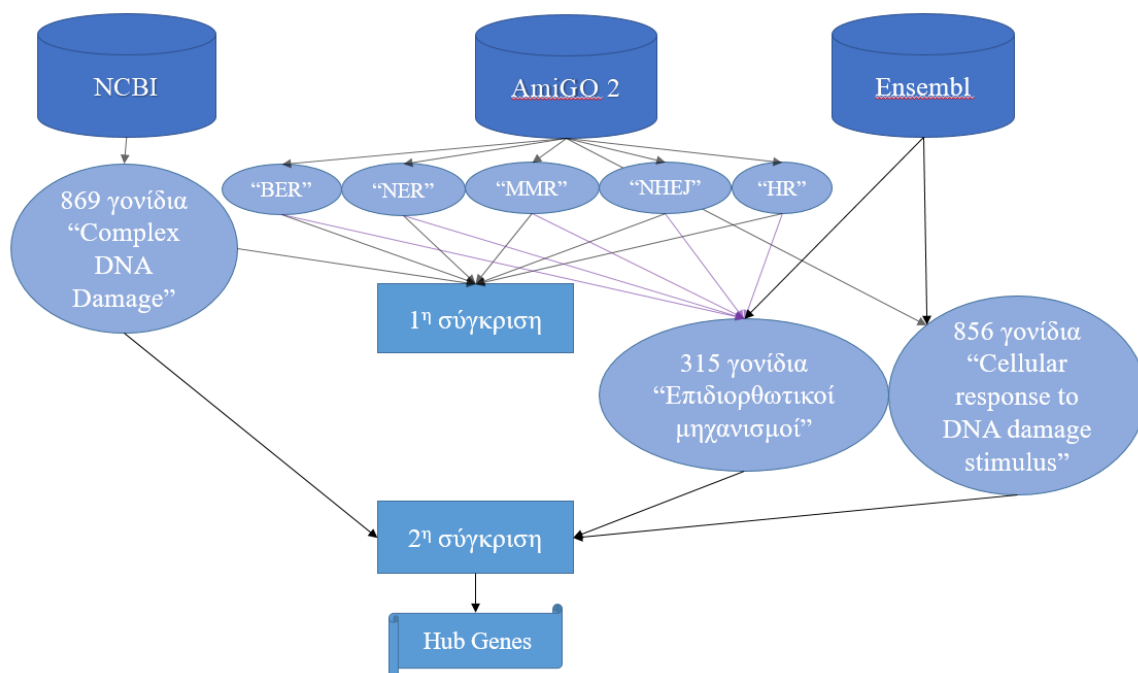
Κεφάλαιο 5: Εύρεση γονιδίων που συμμετέχουν στις σύνθετες βλάβες του DNA

5.1 Σκοπός

Ο πρώτος στόχος της παρούσας διπλωματικής εργασίας είναι ο εντοπισμός των γονιδίων που σχετίζονται με την απόκριση των ανθρώπινων κυττάρων στην ύπαρξη σύνθετων βλαβών στο γενετικό τους υλικό.

Επιπλέον είναι απαραίτητο αυτό το σύνολο γονιδίων να φιλτραρισθεί, ώστε για την περαιτέρω ανάλυση να μην ασχοληθούμε με γονίδια που είναι ήδη γνωστά στην επιστημονική κοινότητα για τον ρόλο τους και τις βιολογικές τους λειτουργίες αλλά και να απαλλαγούμε από τυχόν γονίδια που από λάθος των αλγορίθμων αναζήτησης βρέθηκαν στα αποτελέσματα μας. Σε πολλές μελέτες και ερευνητικές εργασίες αναφέρονται γονίδια που φαίνονται να ενεργοποιούνται όταν δημιουργούνται σύνθετες βλάβες στο DNA και τα οποία δεν ανήκουν σε κάποιον από τους γνωστούς όρους γονιδιακής οντολογίας, αφού δεν υπάρχει ακόμα όρος «απόκριση σε σύνθετες βλάβες DNA» [49, 50]. Αυτός είναι ο λόγος για τον οποίο απαιτείται το φιλτράρισμα και η απομάκρυνση όσων γονιδίων είναι ήδη γνωστή η λειτουργία τους, ώστε να αναλύσουμε τα εναπομείναντα για τα οποία δεν υπάρχουν πολλές πληροφορίες. Για την πραγματοποίηση του φιλτραρίσματος χρησιμοποιήθηκαν τα γονίδια που είναι γνωστό ότι ανήκουν στους επιδιορθωτικούς μηχανισμούς του DNA και επίσης τα γονίδια που ανήκουν στον όρο γονιδιακής οντολογίας: «κυτταρική απόκριση σε βλάβες DNA» (cellular response to DNA damage stimulus). Ο συγκεκριμένος όρος δεν περιέχει την λέξη «σύνθετη» που είναι και η κρίσιμη διαφορά με τα γονίδια που μας ενδιαφέρουν.

5.2 Μεθοδολογία



Διάγραμμα 5.1: Συνοπτική παρουσίαση της μεθόδου που ακολουθήθηκε. Οι κύλινδροι αναπαριστούν τις βάσεις δεδομένων ενώ οι ελλείψεις τα σύνολα των γονιδίων

Η αναζήτηση για τα γονίδια που σχετίζονται με την απόκριση στις σύνθετες βλάβες του DNA έγινε στην ιστοσελίδα της NCBI και συγκεκριμένα στην βάση δεδομένων Gene, η οποία όπως φαίνεται και από την ονομασία της περιέχει όλα τα γνωστά γονίδια. Το λήμμα αναζήτησης που χρησιμοποιήσαμε και στο οποίο καταλήξαμε δοκιμάζοντας τα επιμέρους λήμματα που αποτελούν το τελικό, είναι το εξής: “Complex DNA damage” OR “Clustered DNA damage” OR “Complex DNA lesions” OR “Clustered DNA lesions” AND “homo sapiens”. Χρησιμοποιώντας δηλαδή όλους τους όρους που έχουν χρησιμοποιηθεί στην βιβλιογραφία για να περιγράψουν τις σύνθετες βλάβες στο DNA και με την επιλογή OR ανάμεσα τους είμαστε βέβαιοι ότι δεν θα υπάρξει κάποιο γονίδιο από όσα ενεργοποιούνται με την ύπαρξη σύνθετων βλαβών στο DNA που δεν θα συμπεριληφθεί στα αποτελέσματα μας. Επίσης ο προφανής όρος homo sapiens και η επιλογή “AND” περιορίζει τα αποτελέσματα μόνο σε όσα έχουν να κάνουν με

τον άνθρωπο και όχι οποιονδήποτε άλλο οργανισμό. Το αποτέλεσμα της αναζήτησης είναι 869 γονίδια που απεικονίζονται στον πίνακα 5.1 της επόμενης ενότητας.

Στην συνέχεια αναζητήθηκαν τα γονίδια που ανήκουν στους πέντε γνωστούς επιδιορθωτικούς μηχανισμούς του DNA ώστε να γίνει μια πρώτη σύγκριση με το βασικό σύνολο γονιδίων και να φανεί πόσα από τα γονίδια της απόκρισης στις σύνθετες βλάβες DNA είναι ήδη γνωστό ότι συμμετέχουν στους επιδιορθωτικούς μηχανισμούς του DNA. Η αναζήτηση έγινε στην ιστοσελίδα AmiGO2 με βάση τους όρους γονιδιακής οντολογίας που είχε χρησιμοποιήσει η Δρ. Νικητάκη στην διδακτορική της διατριβή. Οι όροι είναι οι εξής (σε παρένθεση αναγράφεται ο μηχανισμός στον οποίο αντιστοιχούν): GO:0006284 (BER), GO:0006289 (NER), GO:0006298 (MMR), GO:0000724 (HR), GO:0006303 (NHEJ). Με το επιπλέον φίλτρο αναζήτησης “*homo sapiens*” ώστε τα αποτελέσματα να αφορούν μόνο ανθρώπινα γονίδια, συγκεντρώθηκαν τα γονίδια που συμμετέχουν στους επιδιορθωτικούς μηχανισμούς (βλέπε πίνακα 5.2) και πραγματοποιήθηκε μια πρώτη σύγκριση με το βασικό σύνολο μέσω διαγράμματος Venn από την ιστοσελίδα Bioinformatics and Evolutionary Genomics (εικόνα 5.1).

Επίσης στην ιστοσελίδα AmiGO2 αναζητήθηκε ο όρος γονιδιακής οντολογίας “*Cellular response to DNA damage stimulus*” (GO:0006974) για να βοηθήσει και αυτός στο φιλτράρισμα των αποτελεσμάτων. Ο ίδιος όρος αναζητήθηκε και στην βάση δεδομένων Ensembl, ενώ επίσης στην Ensembl αναζητήθηκαν και τα γονίδια των επιδιορθωτικών μηχανισμών όπως έγινε ήδη στην AmiGO2, ώστε να υπάρχει η βεβαιότητα ότι δεν χάθηκε καθόλου πληροφορία που ενδεχομένως να υπάρχει μόνο στην μια βάση δεδομένων και όχι στην άλλη. Τέλος τα γονίδια που συμμετέχουν στους επιδιορθωτικούς μηχανισμούς συγκεντρώθηκαν σε ένα ενιαίο σύνολο με ονομασία «επιδιορθωτικοί μηχανισμοί DNA». Καταλήγουμε με 856 γονίδια που ανήκουν στον όρο “*Cellular response to DNA damage stimulus*” και 355 γονίδια που αποτελούν τους επιδιορθωτικούς μηχανισμούς (αφού βέβαια αφαιρεθούν τα διπλότυπα γονίδια που συμμετέχουν σε παραπάνω από έναν επιδιορθωτικούς μηχανισμούς ώστε να υπάρχει μία φορά το κάθε γονίδιο).

Επόμενο βήμα είναι η εύρεση κοινής βάσης ονοματολογίας για τα γονίδια ώστε να είναι δυνατή η σύγκριση τους. Οι βάσεις δεδομένων NCBI και AmiGO2 δίνουν τα γονίδια με την επίσημη ονομασία τους (official gene symbol) ενώ η Ensembl έχει ξεχωριστό σύστημα ονοματολογίας, τα ensembl ID. Η σύγκριση τελικά γίνεται με βάση την επίσημη ονομασία των γονιδίων. Στην ειδική πλατφόρμα που υπάρχει στην ιστοσελίδα της Ensembl τοποθετούνται οι λίστες με τα γονίδια (όσα έχουν ονομασία κατά ensembl) και επιλέγεται μετατροπή από ensembl ID σε official gene symbol. Κατά την μετατροπή από ensembl ID στην επίσημη ονομασία παρατηρήθηκε ότι για τους επιδιορθωτικούς μηχανισμούς ο αριθμός των γονιδίων μειώνεται, κάτι που δικαιολογείται με το ότι όταν μια ερευνητική ομάδα ανακαλύπτει ένα γονίδιο του δίνει ένα καινούργιο και μοναδικό ensembl ID μέχρι να προσδιορισθεί αν όντως πρόκειται για νέο γονίδιο ή αν ταυτίζεται με κάποιο ήδη γνωστό και επομένως το επίσημο όνομα του δεν διαφέρει. Τα τελικά σύνολα είναι τα:

Σύνολο A → επιδιορθωτικοί μηχανισμοί DNA (315 γονίδια) (πίνακας 5.3)

Σύνολο B → cellular response to DNA damage stimulus (856 γονίδια) (πίνακας 5.4)

Σύνολο C → complex DNA damage (869 γονίδια) (πίνακας 5.1)

τα οποία και συγκρίνουμε μέσω διαγράμματος Venn (εικόνα 5.2).

Το τελευταίο βήμα για το φιλτράρισμα των αποτελεσμάτων είναι ο προσδιορισμός των hub genes του συνόλου με τα γονίδια των Complex DNA damage. Τα hub genes ορίζονται ως τα γονίδια με την μεγαλύτερη συσχέτιση - συνδεσιμότητα με τα υπόλοιπα γονίδια ενός συνόλου [51]. Δηλαδή σε ένα σύνολο γονιδίων, εκείνα που αλληλεπιδρούν ή/και συσχετίζονται με τα περισσότερα γονίδια αποτελούν τα hub genes. Είναι κατά μια έννοια τα πιο σημαντικά γονίδια του συνόλου. Για την συγκεκριμένη εργασία θεωρήσαμε ως hub genes όσα γονίδια εμφανίζουν 5 ή περισσότερες αλληλεπιδράσεις. Τα hub genes προσδιορίζονται με την βοήθεια του εργαλείου STRING που εντοπίζει τις αλληλεπιδράσεις μεταξύ γονιδίων. Σαν input

στο συγκεκριμένο εργαλείο δίνουμε τα 869 γονίδια του συνόλου C και επιλέγουμε σαν οργανισμό τον *homo sapiens*. Ως φίλτρα επιλέγουμε: Active interaction sources → Experiments και Minimum required interaction score → High confidence δηλαδή για τις αλληλεπιδράσεις επιλέγονται εκείνες που έχουν επιβεβαιωθεί πειραματικά και μόνο όσες είναι αρκετά σίγουρο ότι ισχύουν. Οι αλληλεπιδράσεις μεταξύ των γονιδίων απεικονίζονται στην εικόνα 5.3 ενώ το τελικό αποτέλεσμα με τα hub genes του συνόλου C απεικονίζεται στην εικόνα 5.4.

5.3 Αποτελέσματα

Από την αναζήτηση στην NCBI προκύπτουν 869 γονίδια που σχετίζονται με την απόκριση των ανθρώπινων κυττάρων στην ύπαρξη σύνθετων βλαβών στο DNA τους και συγκεντρώνονται στον πίνακα 5.1 παρακάτω.

Αντίστοιχα από την αναζήτηση στην AmiGO2 εντοπίζονται τα γονίδια που συμμετέχουν στους 5 επιδιορθωτικούς μηχανισμούς του DNA τα οποία συγκεντρώθηκαν στον πίνακα 5.2 και κατασκευάστηκε το διάγραμμα Venn για να βρεθεί πόσα από τα γονίδια που σχετίζονται με την απόκριση στις σύνθετες βλάβες DNA είναι γνωστό ότι ανήκουν και στους επιδιορθωτικούς μηχανισμούς. Το διάγραμμα Venn απεικονίζεται στην εικόνα 5.1.

Πίνακας 5.1: Τα 869 γονίδια που συμμετέχουν στην απόκριση των ανθρώπινων κυττάρων στις σύνθετες βλάβες του γενετικού τους υλικού

BRCA1	PML	BIRC3	ITGB3	MCL1	INS	IL2	AGT	PRF1	PTH	DAZL	MX1
ERCC2	ARID1A	SOX9	NUP88	BM1	WNT5A	CFH	BMPR1A	PDGFA	JAK3	ZNF652	PRKDC
TP53	HLA-A	SOX2	CD44	BCL2L1	EGF	MAVS	BMPR2	SFRP2	CSNK1E	HTRA1	CTSS
BRCA2	IGFBP3	TP73	DICER1	BMP2	NOS2	WNT3A	SOST	SOS1	SMC6	ZC3HAV1	DUOX1
BLM	PMS2	RAD50	MMP9	AURKA	VIM	PPP2R1A	TIMP2	HOXB2	CD63	DHFR	CTAG1B
XPC	CD74	NEIL2	H2AFX	SOX17	COP55	IGF2BP3	IL15	AXL	CSF2	CASP7	CXCL13
MSH2	EP300	PTEN	TERC	VHL	MYOCD	PKM	PTOV1	NF1	ENPP1	TREM2	SLIT2
POLG	TP63	SMUG1	NKX3-1	GATA6	RASSF1	XIAP	CCNA1	WEE1	PTX3	TNFSF13	ODC1
PCNA	TREX1	ESR2	GSTP1	SLC6A4	IL17A	PPIA	TET2	DYRK1A	MIR26A1	LDHA	HMBS
RPA1	NFKBIA	ALKBH3	CDKN2A	IL4	HGF	IRS1	ERBB3	PAX1	MIR200B	MIR24-1	PRDX2
APEX1	HDAC1	MAPK1	KRAS	CDC6	TLR9	FOSL2	NGF	TRPV1	MMP12	CSF3	ABCC6
XRCC6	HLA-DQB1	MEN1	CDX2	LEP	MIR155	ABO	BUB1B	GRN	SMC5	UCN	SLC5A5
DDIT3	ERCC8	CUL4A	PTGS2	CTCF	ACE	CCN2	MIR27A	TTN	CXCR3	MUC4	CSPG4
TOP2A	SP1	XRCC1	MBD2	NME1	PDGFRA	IL12B	EDN1	TCEA1	DDX41	MMP8	MUC7
STAT3	HIF1A	POLN	GLI3	KIT	SHC1	TEK	ASXL1	CD40LG	KRT5	TRPV4	NEWENTRY
NFKB1	EGFR	MITF	IRF3	PYCARD	GSTT1	RNASEH2B	S100A4	SERPINB5	NDP	COMP	PIM2
POLD1	LEF1	IL6	ITGAV	ZEB2	BACH2	GNAQ	HUWE1	SPARC	TSLP	ANGPT1	FAP
EZH2	POLI	CCND1	SP1	IRF6	PRMT5	XPO1	TRAF3	LGALS1	TNFSF4	EDNRB	MIR124-2
MRE11	NEIL1	IGF1	PLK1	REL	FOXO1	TXN	MIR29A	EEF1A1	CAT	IL1RL1	FANCB
RAD51	ID3	BCL2	SNAI1	FGF2	BAP1	GJA1	PDCD1	CYP1B1	CADM1	MIR1	IGH
DDB1	MSH6	KMT2A	GLI1	CXCR4	ITGA2	SMAD6	TNFRSF11B	NQO1	TRB	MIR150	TP53INP1
POU5F1	SMARCA4	CRP	RARB	GSTM1	FANCC	FGF	LPL	MPO	CD34	KCNJ5	NLRC4
POLB	MAD2L2	CTCF	PARK7	FOSL1	PDGFB	ANXA2	PROM1	NFATC4	ALCAM	SFTPC	PRTN3
NBN	SMAD3	ERCC6	PTCH1	CDK4	SERPINE1	SIRT6	PIN1	ALDH2	SDC2	HIRT	IL11
TOP1	GATA3	EGR1	SOX10	MIF	FGFR1	BMP4	CDC25A	SF3B4	CRH	PEBP1	CDKN1C
ESR1	AR	ETV6	BHLHE40	BDNF	PLAUR	MIR34A	CXCL10	SELE	ATMIN	HPSE	EVL
SSRP1	POLQ	RBBP8	HMOX1	CASP8	PIK3R1	P2RX7	TGM2	APEX2	CHRNA5	FGF7	IL9
CDKN1A	SMAD4	BAX	BRAF	MMP2	KDR	CYP19A1	BSG	ITGA4	GZMA	EYA4	MIR328
OGG1	NFE2L2	TDP2	FAS	BHLHE41	F3	IFNB1	CDK6	CDK2AP1	IL12A	IGFBP2	FOLR1
CEBPB	CDC73	TRIM28	HR	SATB2	HOXA11	APOA1	LPA	PRAME	STMN1	AFP	TRIM16
MTOR	NTHL1	CDH1	SOX11	NOS3	NOX4	SLC2A1	MIR146A	FSHR	TGFB3	YY1	EPHB6
MGMT	POLL	ITGB1	TEAD1	HSPA1A	VWF	WVVOX	FCGR2B	CTSB	NEFL	MIR141	MIR182
PARP1	TNFAIP3	KLF4	SRC	ABC1	PIK3CG	GBA	ABCG2	CTNBP1	HBEGF	LMAN1	DLL4
CDT1	FOXP3	RBPJ	PAX8	SOD1	FGFR2	YWHAQ	PIM1	TLR3	CD86	HSPH1	LOR
RFC1	RPTOR	PPARD	PIK3CA	SREBF2	ENO1	DAPK1	PLG	TNFSF13B	ZBTB38	WNT1	AGR2
CTNNB1	SIRT1	STAT1	TLR2	KMT2D	FGFR3	KDM5C	ALX4	CRADD	NFIC	CCR7	FST
XPA	TSC2	TP53BP1	CAV1	TFE3	HOXA10	ETV1	SUMO2	IL13	SORT1	SLC11A1	CCL26
MUTYH	CHEK2	TERF1	ADIPOQ	PIAS1	ICAM1	NOTCH2	LCN2	CETP	GLP1R	CSF2RA	RP56KA6
HLA-DRB1	DLCLRE1C	DGCR8	XRCC3	NOD2	LDLR	ABCA1	TFRC	BIN1	NES	CSF2RA	SLC7A1
MYC	MLH1	HES1	HAND2	PINK1	TGFB2	MAP3K7	WRAP53	IL23A	SMOX	TGFB1	SLC2A3
ATM	NOTCH1	MDM2	ERG	COMT	NFIA	CYBA	ACTA2	IFNA1	SORL1	TPH2	DIRAS3
DNMT3A	TSC1	CREBBP	IKZF1	NR0B2	FLT1	TYMP	NTRK2	HOXB4	MAP3K14	CD24	LZTR1
TWIST1	VEGFA	CUL4B	RUNX3	CASP1	PDGFRB	NRAS	DROSHA	MIR125A	MUC5AC	CYP11B2	IHH
DDB2	SMAD2	TLR4	STAT6	AGER	POU4F1	DKK1	NOP53	TNFRSF4	MIR20A	LIPC	UVSSA
PPARG	BCL11B	NR1H3	CLU	MET	ADAR	SKP2	ALOX5	IFIH1	S100A6	MUC2	GPC1
TERF2	DLG1	RUNX1	CXCL8	RIPK1	CD14	IL18	SUMO3	CD38	MIR31	TNKS2	PMEL
HMG2	DTL	IL10	UIMC1	MLH3	CBL	CDKN2B	MAD2L1	DRD1	BGN	TIMP3	MTAP
POLH	IGF1R	CHEK1	FOXO1	SPP1	NOTCH3	CAMP	DOT1L	MC1R	MIR221	ALDH1A1	MMACHC
POLE	TGFB1	APP	SLX4	MKI67	ENG	LEPR	SDC1	PDCD10	RCHY1	LOXL1	NAB2
MCM2	MDC1	NPM1	IRF4	KLF6	RNF168	KPNA2	AGTR1	L1CAM	PRL	PIWIL1	SULT1E1
TARDBP	S100A9	AKT1	UGT1A1	PAX3	BAK1	GJB2	IL1RN	HP	IL22	CASP14	TBXAS1
TDG	SDHB	IL1B	PTPN11	HOXB7	EZR	KLK3	TNFRSF11A	PDCD4	IL27	CA9	PTTG1IP
MCM7	VDR	FAAP24	RAD17	UBE2B	TNFSF11	SOCS3	IFNL1	CD8A	NODAL	NECTIN4	PARP14
TERT	APOBEC3A	PGR	ZEB1	COL1A1	ALB	CAPN3	IL2RA	EPCAM	DLC1	FOLH1	MIR27B
RECQL4	REV1	MICA	ANXA1	FLT3	RAF1	CSF1R	CYP2E1	FLCN	RETN	IL32	MIR181B1
MCM4	WT1	SRF	HEY1	SYK	CD274	VCAM1	AIM2	RNASEH2A	KL	TNFRSF10A	SCGB3A1
POLK	APOE	IGF2	NLRP3	TRERF1	TINF2	PLAU	MYD88	FABP4	CCM2	IL17F	ARL6IP5
GNAS	APC	BIRC5	SRY	TNFRSF1A	RNF8	CYP1A1	CYP2D6	TLE3	LIMS1	MIR96	HSPB6
NFATC1	TAP1	SNCA	SMAD7	MAPK8	RYR1	SF3B1	INHBA	FGF10	GZMB	KRT14	EBP
ERCC1	MAPT	MUC1	CSNK2A1	PON1	POSTN	TNKS	PDIA3	ANGPT2	GCLC	CX3CR1	MAGEC2
ID1	CD247	POLDIP2	AICDA	CLSPN	MMP1	PPM1D	ATAD5	Sep-02	CA2	DMBT1	SMCHD1
HLA-G	FMR1	SMARCA2	MIR21	C9orf72	PTK2	FASLG	SIM1	CD163	CLDN1	P2RY12	RNASEH2C
TNF	NEIL3	RAD18	PROX1	SAMHD1	NHEJ1	ACVR1	GRIN2B	FCGR3A	DEFB4A	SLC6A2	CLIC5
JUN	FOS	MTHFR	RET	ALK	C3	BMX	NMCKS1	FGF1	CTSL	TRAIIP	MMP13
MPG	NKX2-1	NR4A1	CCNB1	ACVRL1	CDC20	MBL2	FKBP5	PITX1	CNTNAP2	HSD11B1	CCR6
HLA-B	APOBEC3B	YAP1	CXCL12	CFL1	ELAVL1	SOD2	CGAS	TNFSF10	IL7	PPP1R1B	LOXL2
MTA2	HSF1	HNF1B	FOXO2	IKBK	DLX4	FHIT	APOB	CNR1	LIF	CCL21	MED12
PRKAR1A	ATR	RAD51B	STAT5A	CCR5	HSPA4	KRIT1	LIN28A	HTT	HSD11B2	COL18A1	NSMCE1
LIG1	RXRA	SREBF1	LGALS3	HRAS	SCN5A	UBE2C	F2R	CD209	DEFB1	MIR106B	LOX
DDX11	RANBP2	ERBB2	PHB	CCL2	IRAK1	COL2A1	FASN	PF4	MMP7	NFATC2	FLG
RB1	HMGA1	IFNG	CD4	GFAP	PRKCZ	OLR1	CTNNA1	TRIB3	TMSB4X	LGR5	GATA2
ID2	GATA4	CDKN1B	HMGA2	CCNA2	LRP5	FGFR4	TIMP1	EGLN3	IDH1	TNC	MIR210
TFF2	AREG	PAX2	CALR								DUSP1

Πίνακας 5.2: Τα γονίδια που συμμετέχουν στους επιδιορθωτικούς μηχανισμούς του DNA σύμφωνα με την βάση δεδομένων AmiGO2.

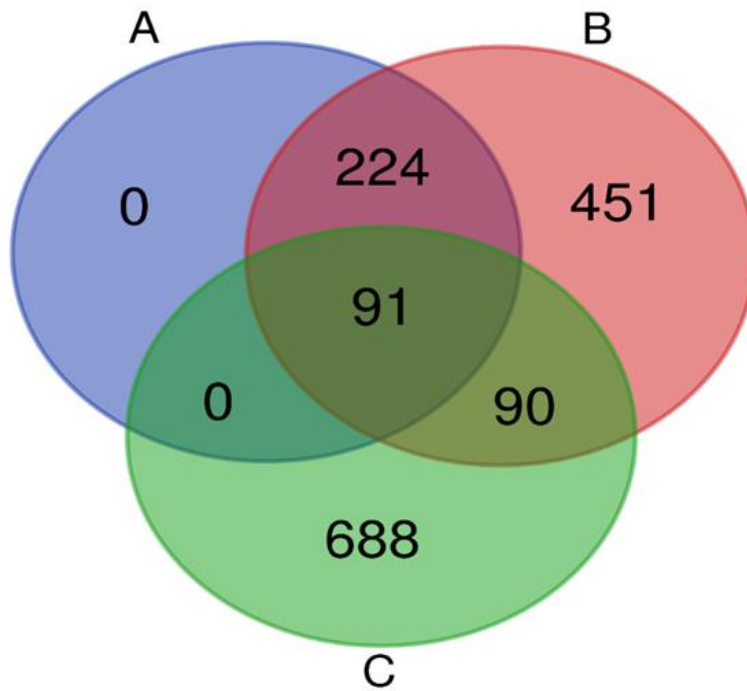
BER	NER		MMR	HR		NHEJ	
PRMT6	COPS7B	POLL	HMGB1	AP5S1	MRNIP	NHEJ1	NSD2
DNA2	LIG4	MMS19	HDAC10	NSMCE2	TEX15	MRE11	RNF8
RPA4	AQR	PNKP	PRKCG	MRE11	PARP1	LIG4	KAT5
ERCC6	PPIE	XPC	PMS2P1	KLHL15	SLX1A	NUDT16L1	BARD1
LIG1	COPS3	ERCC8	MCM9	KDM1A	RAD51	XRCC4	BRCA1
USP47	ERCC3	PARP1	POLD2	GINS4	RAD21L1	ATM	SETMAR
E9PQ18	POLR2F	DBB2	MLH1	RMI1	RPA1	H2AFX	WRAP53
TP53	COPS2	BIVM-ERCC5	RPA2	GEN1	RAD50	PAXX	POLM
HMGB1	BRCA2	RPA1	AXIN2	SEM1	ERCC4	UIMC1	UBE2N
NEIL2	POLD3	POLD1	PMS2P5	FIGL1	CDC7	TFIP11	NBN
FAM168A	UBB	ERCC4	PMS2P3	BRCA2	ZNF365	SHLD2	PRKDC
XPA	CETN2	ERCC2	MSH4	H2AFX	RNF138	XRCC6	ercc1-xpf_human
RPS3	RPA4	MNAT1	POLD3	RPA4	NSMCE1	PRPF19	ku70_80_human
UNG	SLC30A9	CCNH	LIG1	MMS22L	RBBP8	RIF1	ZBTB7A
WRN	ERCC6	RBBP8	PMS2P11	SHLD2	PSMD14	AUNIP	MAD2L2
XRCC1	LIG1	COP56	SETD2	WDR48	SHLD1	ERCC1	KDM2A
SIRT6	POLR2A	RPS27A	TP73	SETD2	PALB2	SHLD3	MDC1
HUWE1	CUL4B	COP58	POLD1	ZFYVE26	RECQL	TP53BP1	MLH1
APEX2	CUL4A	POLR2I	MSH2	INO80	SLX4	PIAS4	LIG3
NEIL1	GTF2H1	COP55	RPA4	EXD2	SMC5	SUMO1	PARP3
POLQ	ISY1	SLX4	TREX1	RIF1	SFR1	BRCC3	HERC2
POLG	GPS1	GTF2H4	ERCC1	AUNIP	HELQ	HSF1	HIST3H3
POLL	PRPF19	RFC1	EXO1	SHLD3	BRCA1	BABAM1	POLL
PARP2	POLR2E	RFC2	PCNA	NABP2	RPA3	SMCHD1	HMGA2
HMGA2	ERCC1	RFC4	RPA1	RFWD3	RAD54L	KDM4D	DNMT
MPG	ZNF830	RPA3	MLH3	MAGEF1	WRAP53	DEK	BABAM2
MUTYH	RAD23B	POLE	PMS2P2	SWSAP1	UBE2N	DTX3L	DCLRE1C
APEX1	GTF2H2	POLD4	PMS2	TP53BP1	NIPBL	PARP9	MRNIP
RPA1	TP53	XAB2	PMS1	FIGL2	NBN	XRCC5	PAXIP1
POLD1	GTF2H5	HUS1	MSH6	MUS81	HUS1	USP51	POLA1
POLB	HMGB1	OGG1	ABL1	RECQL5	RADX	XRCC1	RAD50
SMUG1	NEIL2	POLR2D	TDG	MEIOB	inip-ints3-nabp2_human	CYREN	ERCC4
TDG	CETN1	NEIL3	RPA3	POLN	btr_human	DCLRE1A	ABRAXAS1
RPA3	GTF2H3	NTHL1	POLD4	RTEL1-TNFRSF6B	nua4_human	BLM	UBE2V2
POLE	POLR2K	ercc1-xpf_human	XPC	PARBP	inip-ints3-nabp1_human	UVRAG	POLB
OGG1	XPA	POLR2G	msh2-msh3_human	SMCHD1	MORF4L1	POLQ	ATP23
NEIL3	CHD1L	POLK	RNASEH2A	MCM9	GINS2	HIST1H4A	PSMD14
NTHL1	UBC	POLR2J	MSH3	KDM4D	FEN1	DCLRE1B	RNF168
Q6ZNB5	HMGN1	POLR2H	MSH5	MCMD2C	RAD54B		SHLD1
FEN1	USP7	RNF111	MUTYH	REC8	DMC1		
MOR2N6	RAD23A	RFC3		YY1	SAMHD1		
RPA2	POLR2C	RFC5		WRN	MAD2L2		
HMGA1	CDK7	RPA2		SPIDR	LOC107984078		
RECQL4	GTF2H2C	FAN1		RAD52	NABP1		
MBD4	EP300	LIG3		SFPQ	RPA2		
LIG3	UVSSA	POLR2B		USP51	AP5Z1		
	RAD52	COPS7A		SWI5	FAN1		
	FANCC	PCNA		CHEK1	RECQL4		
	USP45	DBB1		XRCC1	TONSL		
	ERCC5	NEIL1		RTEL1	FANCB		
	TCEA1	HUS1B		SIRT6	LIG3		
	COPS4	RBX1		CCDC155	FUS		
	XRCC1	POLR2L		XRCC3	CGAS		
	BRIP1	POLD2		XRCC2	ACTR2		
	UBA52	SIRT1		WAS	POLL		
		ATXN3		BLM	PPP4C		
				NUCKS1	MCM8		
				HUS1B	ZSWIM7		
				CDC45	RAD51AP1		
				POLQ	TERF2IP		
				RAD51C	MMS19		
				RMI2	TIMELESS		
				FBH1	RAD51B		
				FIGN	PPP4R2		
				RAD51D	SMC6		
				HELB			

Πίνακας 5.3: Το σύνολο A. Η πλήρης λίστα των γονιδίων που συμμετέχουν στους 5 επιδιορθωτικούς μηχανισμούς του DNA με την επίσημη ονομασία τους (official gene symbol)

PRMT6	RECQL5	CDK7	KDM1A	KAT5	TREX1
DNA2	MEIOB	GTF2H2C	GIN54	BARD1	ERCC1
USP47	POLN	EP300	RMI1	SETMAR	EXO1
E9PQ18	RTEL1-TNFRSF6B	UVSSA	GEN1	POLM	PCNA
FAM168A	PARPBP	FANCC	SEM1	PRKDC	MLH3
RPS3	SMCHD1	USP45	FIGNL1	ku70_80_human	PMS2P2
UNG	MCM9	ERCC5	BRCA2	ZBTB7A	PMS2
HUWE1	KDM4D	TCEA1	H2AFX	KDM2A	PMS1
APEX2	MCMDC2	COPS4	RPA4	MDC1	MSH6
POLG	REC8	BRIP1	MMS22L	PARP3	ABL1
PARP2	YY1	POLR2B	SHLD2	COPS7B	TDG
MPG	WRN	COPS7A	WDR48	LIG4	POLD4
APEX1	SPIDR	DDB1	SETD2	AQR	XPC
SMUG1	RAD52	NEIL1	ZFYVE26	PPIE	RNASEH2A
Q6ZNB5	SFPQ	RBX1	INO80	COPS3	MSH3
MOR2N6	USP51	POLR2L	EXD2	ERCC3	MSH5
HMGA1	SWI5	SIRT1	RIF1	POLR2F	MUTYH
MBD4	CHEK1	ATXN3	AUNIP	COPS2	AP5S1
NHEJ1	XRCC1	UBA52	SAMHD1	UBB	NSMCE2
NUDT16L1	RTEL1	PNKP	MAD2L2	CETN2	MRE11
XRCC4	SIRT6	ERCC8	LOC107984078	SLC30A9	KLHL15
ATM	CCDC155	DDB2	NABP1	ERCC6	inip-ints3-nabp2_human
PAXX	XRCC3	BIVM-ERCC5	RPA2	POLR2A	btr_human
UIMC1	XRCC2	ERCC2	AP5Z1	ERCC4	nua4_human
TFIP11	WAS	MNAT1	FAN1	CDC7	inip-ints3-nabp1_human
XRCC6	BLM	CCNH	RECQL4	ZNF365	MORF4L1
PIAS4	NUCKS1	COPS6	TONSL	RNF138	GIN52
SUMO1	HUS1B	RPS27A	FANCB	NSMCE1	FEN1
BRCC3	CDC45	COPS8	LIG3	RBBP8	RAD54B
HSF1	POLQ	POLR2I	FUS	PSMD14	DMC1
BABAM1	RAD51C	COPS5	CGAS	SHLD1	CUL4B
DEK	RMI2	GTF2H4	TIMELESS	PALB2	CUL4A
DTX3L	FBH1	RFC1	RFWD3	RECQL	GTF2H1
PARP9	FIGN	RFC2	MAGEF1	SLX4	ISY1
XRCC5	RAD51D	RFC4	SWSAP1	SMC5	GPS1
CYREN	HELB	POLE	TP53BP1	SFR1	PRPF19
DCLRE1A	POLL	XAB2	GTF2H2	HELQ	PMS2P11
UVRAG	PPP4C	OGG1	TP53	BRCA1	TP73
HIST1H4A	MCM8	POLR2D	GTF2H5	RPA3	POLD1
HERC2	ZSWIM7	NEIL3	NEIL2	RAD54L	MSH2
HIST3H3	RAD51AP1	NTHL1	CETN1	WRAP53	NABP2
HMGA2	TERF2IP	ercc1-xpf_human	GTF2H3	UBE2N	FIGNL2
DNTT	MMS19	POLR2G	POLR2K	NIPBL	msh2-msh3_human
BABAM2	ACTR2	POLK	XPA	NBN	POLR2E
DCLRE1C	RAD51B	POLR2J	CHD1L	HUS1	PMS2P1
PAXIP1	PPP4R2	POLR2H	UBC	RADX	POLD2
POLA1	SMC6	RNF111	HMGN1	AXIN2	MLH1
ABRAXAS1	MRNIP	RFC3	USP7	PMS2P5	RAD21L1
UBE2V2	TEX15	RFC5	RAD23A	PMS2P3	RPA1
POLB	PARP1	HMGB1	POLR2C	MSH4	RAD50
ATP23	SLX1A	HDAC10	MUS81	POLD3	RAD23B
RNF168	RAD51	PRKCG	SHLD3	LIG1	ZNF830
DCLRE1B	NSD2	RNF8			

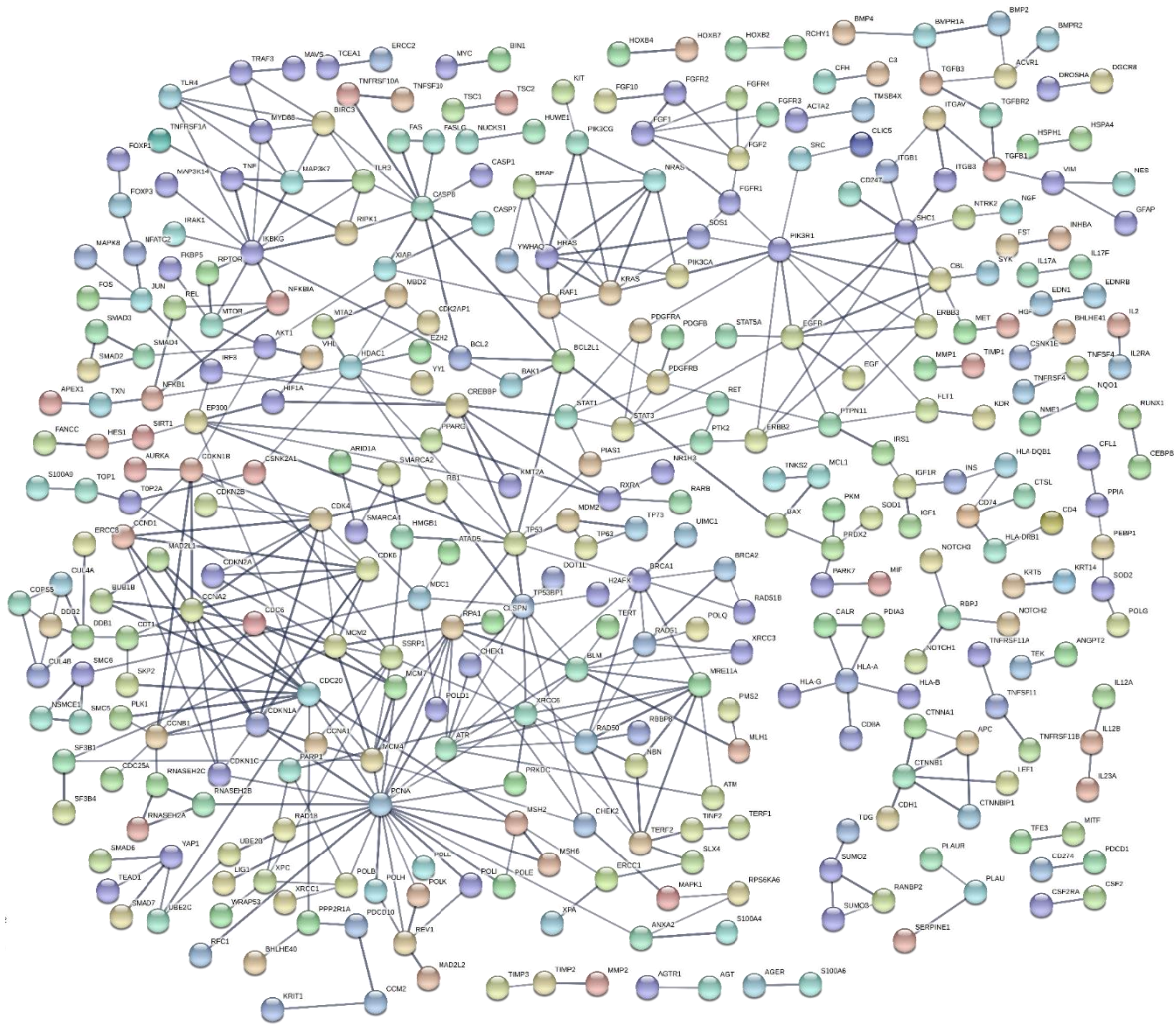
Πίνακας 5.4: Τα 856 γονίδια που ανήκουν στον όρο “Cellular response to DNA damage stimulus”

POLN	Q6ZNB5	TEX12	BCL2L1	CNOT2	GTF2H1	CNOT3	BLM	HINFP	TANK	MORC2	SP100
UBE2L6	RNF111	TEX15	CBX3	NOP53	MCM7	CNOT4	UBE2A	POLR2E	PMS2P1	CDK1	CCDC155
TP53BP1	FEN1	PAXIP1	BAX	UBE2V1	PLAGL1	AP5Z1	NUCKS1	TREX1	XPA	RPS27A	CLU
FIGNL2	TOPORS	A0A0U1RC	PAGR1	PMS2P3	MYO6	NABP1	NACC2	ERCC1	ZBTB18	COPS8	WDHD1
MRPS26	NSMCE3	POLA1	INO80D	BAG6	ATMIN	TMEM189	APEX1	EXO1	CBX8	RAD9A	ASH2L
HMGB1	NSMCE2	PARP1	ELL3	RBM24	ISY1	AXIN2	RPA1	CEP164	SPDYA	SETD7	RBM17
SLF1	MRE11	ATM	FANCE	SMARCAL1	FZR1	ACER2	SPRTN	LYN	MAPK12	BAD	CFAP410
MUS81	POLE3	TIPIN	RAD18	C9orf116	ZBTB42	FAN1	NEIL1	FBXO6	SUMO1	AKT1	POLH
ZDHHHC16	LIG4	CINP	CEP63	BRAT1	CSNK1E	MASTL	UVRAG	SHLD3	APTX	RNF8	POLR2B
ZRANB3	NUPR1	UHRF1	INTS7	RPL26	CNOT6	KDM2A	CDCA5	ZNF830	CHD1L	ATRIP	RUVBL1
POLE2	ACD	TERF2IP	MACROD1	POLR2F	INO80	ZBTB1	DDX11	TREX2	TAOK1	CCNB1	TAF9
HDAC10	KLHL15	MMS19	URS00001	COPS2	PDS5B	USP10	CASP2	RAD23B	ASCC1	SMUG1	COPS7A
NEIL2	KDM1A	ACTR2	TLK1	MSH4	MCT51	RECQL4	PTTG1	PML	RTEL1-TNF	SUPT16H	ALKBH5
BACH1	PRIMPOL	PNKP	RAD51D	FIGNL1	PPP1R15A	NCOA6	ALKBH8	NABP2	PARBPB	RECQL	APEX2
CD44	FMR1	HMOX1	UBA1	DONSON	HERC2	SMC1A	POLG2	UNG	TMEM161	SGK1	PCNA
CETN1	AQR	AURKA	CDC14B	MTA1	PMS2P2	MDC1	HUS1B	FOXN3	BABAM1	CETN2	BCL2L11
FOXO1	FOXM1	FANCA	SMG1	APBB1	PIK3R1	ZC3H12A	CNOT8	XRCC5	JMY	RNF61	FOXO3
MARH7	UBE2E2	TMEM109	RHNO1	SLF2	CNOT11	TRIP12	MCRS1	CHCHD4	HPF1	GTF2H2	BCL2A1
CDC25C	TNKS1BP1	TP73	ZBTB5	BRCA2	HIST3H3	IRF3	MAPT	EMSY	WRNIP1	SPATA18	MAPK14
UBE2T	FTO	CDKN2AIP	ZNF385A	DYRK2	FBXO4	TONSL	SYF2	MGMT	NEDD4	APLF	DDB1
ISG15	CNOT6L	ETAA1	RCHY1	H2AFX	IFI16	ASCC2	SPATA22	SFPQ	E2F8	TP53	XRCC3
DTL	PIF1	AATF	TDP1	CARM1	PMS2	RNF169	CIDEB	FANCC	SMCHD1	TOPBP1	XRCC2
RRM2B	KMT5A	RAD51B	RBM14	DNA2	PMS1	PYHIN1	MLH3	ASTE1	ENDOV	TAF9B	CRY1
REV3L	CSB-PGBD	GADD45A	NHEJ1	MAPK15	ATXN3	SFRP2	TIPRL	CDK3	MCM9	WNT1	BATF
PRKCG	PPIE	DNTT	COPS7B	TFDP3	BAK1	RPS6KA6	CDC45	THOC5	EIF2AK4	DDX39B	WAS
HMG1	ZBTB3	TRIM28	AP5S1	POLD3	SUV39H1	REV1	POLQ	PWWP3A	MDM2	PARG	PARK7
USP7	INO80B	BABAM2	PEA15	MORF4L2	HTRA2	PSEN1	FBXO45	USP43	NSMCE4A	GTF2H5	MNDA
RAD23A	PRMT6	INO80C	USP1	PCLAF	PARP4	PARP10	DMAP1	USP45	IMMP2L	MAGEF1	ACKR3
MRPS35	CBL	PPP4R2	CXCL12	UBB	RAD51C	PRMT1	ZFYVE26	USP51	ZBTB4	A8MVJ9	CHEK1
POLR2C	NIPBL	MSX1	POLD4	EYA2	MSH6	CBX1	PLK3	ERCC5	KDM4D	CLSPN	CRIP1
ATF2	UBE2D3	POLD1	MEN1	UBA7	EPC2	AEN	MAEL	SWI5	BCCIP	DDX1	TPT1
CDK9	NBN	EYA1	XAB2	ESCO2	CASP3	ARID3A	RAD17	VCP	FNIP2	PMAIP1	NEK1
POLI	HUS1	MYC	RAD54L	PRKDC	GRB2	CIB1	DDX5	CHRNA4	CGAS	STXBP4	APC
REC8	USP28	PRKCD	NUDT1	SHLD2	UBA52	ZNF365	USP47	EME1	MCMDC2	MRNIP	COPS4
UBR5	PAXX	BCL2L2	BCL2L10	SLC30A9	E2F4	TERF2	YJU2	NUAK1	TIMELESS	FANCD2	TCEA1
TRIM25	PPM1D	MSH2	MEIOC	SESN2	HELB	RNF138	NEK11	MPG	CNOT1	NTHL1	FEM1B
ING4	FANCG	TOP2A	TNF	WDR48	POLL	POLB	FIGN	DDB2	CBX5	BTG2	BCL2L12
IER3	UIMC1	RAD50	DDIAS	ERCC6	PARP2	PMS2P5	POLD2	STUB1	WDR33	ASF1A	IGHMMP2
ATR	REXO4	TRAF6	NPM1	TLK2	PPP4C	SIRT4	PSMD10	YY1	EPHA2	MORF4L1	SHPRH
RBBP6	TFIP11	ERCC4	CEBPG	RINT1	USP16	NEK4	BCL2	RAD21	CDK7	POLR2G	RBX1
SNW1	RPA4	ERCC2	TAF1	RPAIN	SNAI1	ATP23	ZMPSTE24	SLX1A	WRN	RNASEH2A	POLR2L
NUPR2	PAK1	FAAP100	DCLRE1C	ZMAT3	MAPKAPK2	NPAS2	ZBTB47	UBE2NL	GTF2H2C	POLK	NUDT16
RECQL5	CHAF1B	ABRAXAS1	XPC	CDK2	DYRK3	MNAT1	NFATC4	RGCC	TFDP2	MRPS9	RM12
BCL6	CHAF1A	UBE2V2	ERCC8	LIG1	MCM8	CENH	MAP2K6	RAD51	TFDP1	TFPT	POLG
CDC7	BRD4	MCM10	SMC6	ALKBH1	UBE2U	NSMCE1	HIPK2	TIGAR	MICA	CENPX	HIST1H4A
MEIOB	MIF	RAD9B	WRAP53	POLR2A	TRIAP1	ANKRD1	FMN2	BOK	SMYD2	RAD54B	FBH1
SPRED1	MMS22L	GINS4	POLM	PMS2P11	HMG2	PSMD14	NFRKB	E2F1	SPIDR	URS00003	PARP3
SPRED2	FAAP24	ACTR5	ATRX	RBL2	ING2	RNF168	ALKBH2	TNFRSF1B	KIN	EGLN3	UCHL5
CDKN2D	BID	DHX9	UBE2N	DYRK1A	CNOT7	UFD1	FANCB	SMARCAD1	EP300	RFWD3	SPO11
RPS3	TNP1	CHCHD6	ERCC6L2	CUL4B	EME2	COPS6	MBD4	RAD21L1	BIVM-ERC	GIGYF2	PSME4
BRCC3	E9PQ18	RAD1	RNF113A	CUL4A	EGRF	UBE2B	TP63	PALB2	DEK	IKBKE	UBE2W
URS00003	CCND1	RM1	pcna_hum	SETD2	ZSWIM7	USP3	MLH1	SLX4	DDIT3	EXO5	RFC3
CASP9	MDM4	CNOT10	ube2n-ube	CTC1	RAD51AP1	EYA3	ZBTB38	GTF2H4	FUS	INIP	RFC5
UBC	WAC	ZBTB40	inip-ints3-r	METTL3	ABL1	SHLD1	CDK5RAP3	SMC5	CNOT9	FANCF	MOR2N6
HSF1	GPS1	NUDT16L1	ercc1-xpf	STK33	BCL3	HMGB2	LIG3	CDKN1A	HIST3H2A	TRRAP	GINS2
ASCC3	XRCC6	GEN1	eme1-mus	CDKN1B	ZNF771	DCLRE1B	ANKLE1	YAP1	SIX6OS1	KIF22	MRPS11
PIDD1	PRPF19	ARMT1	msh2-msh	RBBP8	BARD1	HIPK1	PYCARD	SFR1	DTX3L	MSH3	RUVBL2
EEDP1	FANCL	ACTR8	btr_humar	IKBKG	DNAJA1	DOT1L	SIRT1	HELQ	BRSK1	POLR2J	TENT4A
GTF2H3	TFAP4	WDR76	rrm1-rrm2	MACROD2	BRCA1	CCDC13	EEF1E1	BCLAF1	TTC5	POLR2H	INO80E
CENPS	TRIM32	SEM1	ku70_80_f	A0A1W2PC	PHLDA3	RBBP5	DDIT4	INTS3	MAPK1	SAMHD1	ZBTB7A
POLR2K	ATAD5	SKIL	nua4_hum	SYCP1	RNASEH2B	FANCM	MAPK3	XRCC1	UVSSA	MAD2L2	PCBP4
PTTG1IP	PPP1R10	FOXO4	ddb1-ddb2	PPP5C	PLK1	NSD2	MUC1	SLFN11	CCAR2	OTUB1	DMC1
MARF1	PLA2R1	XIAP	pp4_huma	POLR2I	RPS27L	ACTL6A	BOD1L1	RTEL1	EID3	TDP2	HIC1
PIAS4	FANCI	COPS3	eme2-mus	KAT5	SPIRE2	CHEK2	RPA2	PLK2	SHISA5	MSH5	MAP3K20
TICRR	EXD2	ZNF652	inip-ints3-r	IRF7	TWIST1	NONO	ADPRHL2	GTSE1	GML	PHF1	TAOK3
MOAP1	ALKBH7	ERCC3	SWSAP1	GNL1	TRIP13	RAD51L3-F	OARD1	CYREN	CDCSL	LOC10798	FBXO5
SUSD6	RIF1	TNFRSF1A	NPLOC4	STK11	SPIRE1	FAAP20	SSRP1	DCLRE1A	RAD52	CDIP1	VAV3
HNRNPK	ALKBH3	FBXO31	NEIL3	CTLA4	NKX3-1	OGG1	EYA4	SIRT6	PARP9	CD74	FOXP1
SETX	SOX4	MGME1	CRADD	COPS5	SETMAR	FGF10	HMGA1	HUWE1	SFN	XRCC4	E2F7
FAM168A	AUNIP	SMC3	GGN	UPF1	TDG	CLOCK	CCNK	BRIP1	CASP14	MC1R	BAZ1B
MCL1	RPA3	TAOK2	PDS5A	NFATC2	RFC1	POLR2D	RFC4	PTPN11	SNAI2	WDR70	MUTYH
					RFC2	RADX				POLE	THOC1



Εικόνα 5.2: Το διάγραμμα Venn των συνόλων A, B και C (A: επιδιορθωτικοί μηχανισμοί DNA, B: cellular response to DNA damage, C: complex DNA damage)

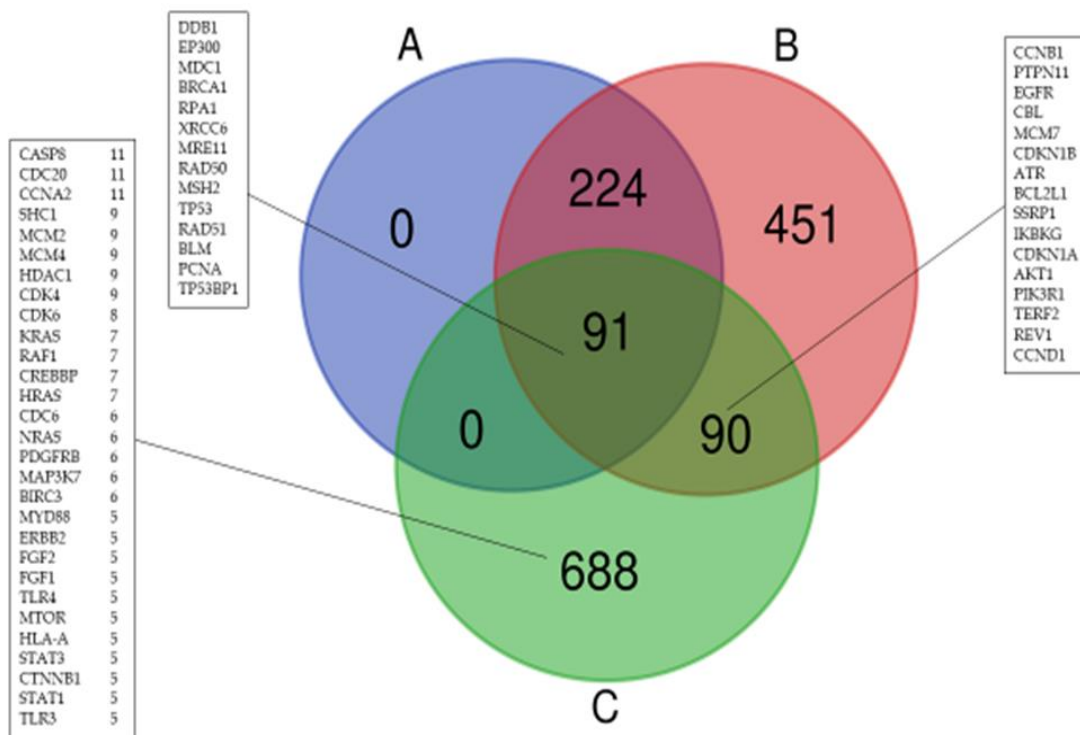
Για το σύνολο C με τα γονίδια που συμμετέχουν στην απόκριση στις σύνθετες βλάβες του γενετικού υλικού εντοπίστηκαν τα hub genes με την βοήθεια της πλατφόρμας STRING. Δίνοντας σαν input τα 869 γονίδια και επιλέγοντας τις κατάλληλες παραμέτρους λαμβάνουμε ως αποτέλεσμα τα protein - protein interactions του συγκεκριμένου συνόλου γονιδίων. Οι αλληλεπιδράσεις απεικονίζονται στην εικόνα 5.3. Όσα γονίδια δεν εμφανίζουν καμία αλληλεπίδραση δεν απεικονίζονται καθόλου.



Εικόνα 5.3: Τα protein - protein interactions του συνόλου C

Όσα γονίδια εμφανίζουν 5 και περισσότερες αλληλεπιδράσεις προκρίνονται ως σημαντικά (hub genes) και έτσι επιτυγχάνεται το φιλτράρισμα των γονιδίων που ήταν και ο αρχικός μας στόχος. Στην εικόνα 5.4 απεικονίζεται το διάγραμμα Venn όπως προέκυψε από την σύγκριση των τριών συνόλων γονιδίων A, B και C (εικόνα 5.2) στο οποίο τώρα φαίνονται και τα hub genes του συνόλου C. Επιπλέον για όσα από τα hub genes ανήκουν αποκλειστικά στο σύνολο C και όχι σε κάποια από τις

τομές του με τα άλλα δυο σύνολα (τα οποία είναι και αυτά που τελικά θα αναλύσουμε περαιτέρω - βλέπε παράγραφο 5.4 για τα συμπεράσματα), αναγράφεται δίπλα και ο αριθμός των αλληλεπιδράσεων που εμφανίζουν όπως προέκυψαν από την πλατφόρμα STRING. Για παράδειγμα το γονίδιο CASP8 εμφανίζει 11 αλληλεπιδράσεις όπως φαίνεται στην εικόνα 5.4. Αυτό σημαίνει ότι αλληλεπιδρά με 11 διαφορετικά γονίδια που ανήκουν στο σύνολο των 869 γονιδίων απόκρισης στις σύνθετες βλάβες DNA (σύνολο C).



Εικόνα 5.4: Τα hub genes του συνόλου C, δηλαδή τα γονίδια που εμφανίζουν 5 ή και περισσότερες αλληλεπιδράσεις με το σύνολο των γονιδίων που ανήκουν στο σύνολο C. Για όσα ανήκουν αποκλειστικά στο σύνολο C αναγράφεται και ο αριθμός των αλληλεπιδράσεων.

5.4 Συνοψιση αποτελεσμάτων - συμπεράσματα

Ο βασικός στόχος στο πρώτο αυτό μέρος της εργασίας ήταν ο εντοπισμός των γονιδίων που σχετίζονται με την απόκριση των ανθρώπινων κυττάρων στην ύπαρξη σύνθετων βλαβών στο γενετικό τους υλικό. Φυσικά μια αναζήτηση στην βάση δεδομένων της AmiGO2 ή της Ensembl δεν θα μπορούσε να είναι αρκετή αφού γνωρίζουμε ότι δεν υπάρχει όρος γονιδιακής οντολογίας «απόκριση στις σύνθετες βλάβες DNA» που να συγκεντρώνει αυτά τα γονίδια. Επομένως αυτή η διαδικασία αναζήτησης στην NCBI ήταν περισσότερο κάτι σαν text mining ώστε να συγκεντρωθούν όσα γονίδια αναφέρονται σε οποιαδήποτε δημοσίευση που αφορά σύνθετες βλάβες DNA. Γνωρίζουμε βέβαια ότι οι αλγόριθμοι αναζήτησης σε αυτές τις βάσεις δεδομένων δεν είναι τόσο εξελιγμένοι επομένως στο σύνολο με τα 869 γονίδια θα υπήρχαν αρκετά γονίδια που από λάθος βρέθηκαν στα αποτελέσματα. Στην πρώτη σύγκριση που κάναμε του βασικού μας συνόλου γονιδίων (869 γονίδια - σύνολο C) με τους επιδιορθωτικούς μηχανισμούς του DNA με το Venn της εικόνας 5.1 ενώ βλέπουμε αρχικά μια συμφωνία με τα αποτελέσματα της διατριβής της Δρ. Νικητάκη που είχε καταλήξει ότι δεν υπάρχει γονίδιο που να συμμετέχει και στους πέντε μηχανισμούς επιδιόρθωσης [52] που σημαίνει ότι τουλάχιστον όσον αφορά τους μηχανισμούς επιδιόρθωσης τα σύνολα γονιδίων είναι εντάξει, εντούτοις παρατηρούμε ότι η πλειοψηφία των γονιδίων του συνόλου C δεν ανήκουν σε κάποιον επιδιορθωτικό μηχανισμό, γεγονός που δεν μας επιτρέπει να κάνουμε σημαντικό φιλτράρισμα των γονιδίων. Στην επόμενη σύγκριση που προστίθεται και το σύνολο (σύνολο B) των γονιδίων του όρου “Cellular response to DNA damage stimulus” (εικόνα 5.2) ενώ βλέπουμε όλα τα γονίδια των επιδιορθωτικών μηχανισμών να ανήκουν και στο νέο σύνολο, όπως αναμέναμε αφού ο όρος “Cellular response to DNA damage stimulus” είναι υπερόνολο (εμπεριέχει) των επιδιορθωτικών μηχανισμών, επομένως αντιλαμβανόμαστε ότι αυτά τα δύο σύνολα γονιδίων είναι σωστά, για ακόμα μια φορά τα περισσότερα γονίδια του complex DNA damage (σύνολο C) δεν εμφανίζουν πολλά κοινά σημεία με τα άλλα δύο σύνολα. Φυσικά δεν περιμέναμε απόλυτη ταύτιση των συνόλων αλλά το γεγονός ότι υπάρχουν τόσα πολλά γονίδια

του συνόλου C που δεν ανήκουν ούτε στον όρο “Cellular response to DNA damage stimulus” ούτε στους επιδιορθωτικούς μηχανισμούς υποδεικνύει ότι πολλά από αυτά δεν είναι πραγματικά γονίδια που συμμετέχουν στην απόκριση στις σύνθετες βλάβες DNA. Επομένως πρέπει να ακολουθηθεί μια διαφορετική μέθοδος φιλτραρίσματος. Έτσι επιλέξαμε την σίγουρη λύση των hub genes, όπου από τα 869 γονίδια θεωρούμε ως σημαντικά μόνο όσα αλληλεπιδρούν με τουλάχιστον 5 από αυτά. Για τα γονίδια που εμφανίζουν τόσες αλληλεπιδράσεις μπορούμε να είμαστε αρκετά σίγουροι ότι πρόκειται για πραγματικά λειτουργικά γονίδια και όχι λάθη των αλγορίθμων αναζήτησης, ειδικά από την στιγμή που οι αλληλεπιδράσεις είναι πειραματικά επιβεβαιωμένες όπως ζητήσαμε από την πλατφόρμα STRING. Επιπλέον για όσα από τα hub genes ανήκουν αποκλειστικά στο σύνολο C (βλέπε εικόνα 5.3), το γεγονός ότι οι αλληλεπιδράσεις είναι αρκετά πιθανό να περιλαμβάνουν και τα γονίδια τα οποία συμμετέχουν στους επιδιορθωτικούς μηχανισμούς ή στον όρο “Cellular response to DNA damage stimulus” (δηλαδή τα γονίδια που ανήκουν στις τομές του C με τα A και B) αυξάνει σημαντικά τις πιθανότητες να είναι πράγματι γονίδια που ενεργοποιούνται με την ύπαρξη σύνθετων βλαβών στο γενετικό υλικό και τα οποία δεν έχουν μελετηθεί επαρκώς για να ενταχθούν σε κάποιον επιδιορθωτικό μηχανισμό ή ακόμα καλύτερα σε έναν νέο όρο γονιδιακής οντολογίας με όνομα «απόκριση στις σύνθετες βλάβες DNA». Με αυτό το σκεπτικό στην επόμενη ενότητα αναλύουμε τα 29 hub genes που ανήκουν αποκλειστικά στο σύνολο C ως προς την σχέση τους με διάφορες ασθένειες και μορφές καρκίνου, τα βιολογικά μονοπάτια στα οποία είναι γνωστό ότι ανήκουν και την κυτταρική γήρανση.

Κεφάλαιο 6: Περαιτέρω διερεύνηση των γονιδίων που επισημάνθηκαν ως κρίσιμα

Σε αυτό το κεφάλαιο πραγματοποιείται μια στοχευμένη ανάλυση των 29 γονιδίων που προκρίθηκαν ως σημαντικά με σκοπό την επαλήθευση της συσχέτισης τους με τις σύνθετες βλάβες DNA. Χρησιμοποιήθηκαν τέσσερις διαφορετικές πλατφόρμες βιοπληροφορικής ανάλυσης οι οποίες κρίθηκαν ότι μπορούν να επαληθεύσουν αυτή την συσχέτιση. Συγκεκριμένα ελέγχθηκαν οι βιολογικοί μηχανισμοί που συμμετέχουν αυτά τα γονίδια, οι ασθένειες με τις οποίες συσχετίζονται, τύποι καρκίνου στους οποίους η αυξημένη έκφραση των γονιδίων επηρεάζει την επιβίωση των ασθενών και η συσχέτιση τους με την κατάσταση της κυτταρικής γήρανσης (senescence).

6.1 Βιολογικοί μηχανισμοί

6.1.1 Στόχος

Για τα 29 γονίδια που προκρίθηκαν στην προηγούμενη παράγραφο ως τα πλέον σημαντικά του συνόλου των γονιδίων που συμμετέχουν στην απόκριση των κυττάρων στην ύπαρξη σύνθετων βλαβών στο DNA θέλουμε να εντοπίσουμε τα «βιολογικά μονοπάτια» (biological pathways) στα οποία συμμετέχουν. Πιο συγκεκριμένα θέλουμε να ελέγξουμε αν συμμετέχουν σε ορισμένους βιολογικές διαδικασίες οι οποίες είναι γνωστό από την θεωρία ότι σχετίζονται με τις σύνθετες βλάβες του γενετικού υλικού όπως για παράδειγμα ο κυτταρικός θάνατος.

6.1.2 Μέθοδος

Μέσω της πλατφόρμας DAVID πραγματοποιήθηκε pathway enrichment analysis για τα 29 hub genes. Στην πλατφόρμα DAVID (βλέπε και παράγραφο 4.4.1) ο χρήστης δίνει μια λίστα γονιδίων την οποία θέλει να επεξεργαστεί και στην συνέχεια επιλέγει αν θέλει για την συγκεκριμένη λίστα να ελέγξει «βιολογικά μονοπάτια», ασθένειες που σχετίζονται με αυτά κ.α. ενώ κάθε μια από τις ενότητες αυτές προσφέρει μια σειρά βάσεων δεδομένων για να επιλέξει ο χρήστης αυτή που προτιμάει για τα αποτελέσματα του (πχ για τους μηχανισμούς έχει να επιλέξει ανάμεσα σε GOTERM, Kegg Pathways κ.α.). Ως input τέθηκε η λίστα με τα 29 γονίδια, για οργανισμό επιλέχθηκε ο Homo Sapiens και για output επιλέχθηκε από τις διαθέσιμες βάσεις δεδομένων από τις οποίες προκύπτουν τα αποτελέσματα την GOTERM_BP_ALL (όπου BP: Biological Process). Σαν επιπλέον φίλτρα τέθηκαν για κάθε μηχανισμό ο αριθμός των counts, δηλαδή ο αριθμός των γονιδίων που συμμετέχουν να είναι τουλάχιστον 4 από τα 29 της λίστας και το p-value (στατιστικό μέγεθος του οποίου η τιμή όσο πιο κοντά είναι στο μηδέν τόσο πιο σίγουρο είναι το αποτέλεσμα) να είναι μικρότερο από 0,01. Από τους βιολογικούς μηχανισμούς που λαμβάνονται σαν αποτέλεσμα από την ανάλυση της πλατφόρμας DAVID ξεχωρίζονται εκείνοι που «μαρτυρούν» την ύπαρξη σύνθετων βλαβών στο DNA.

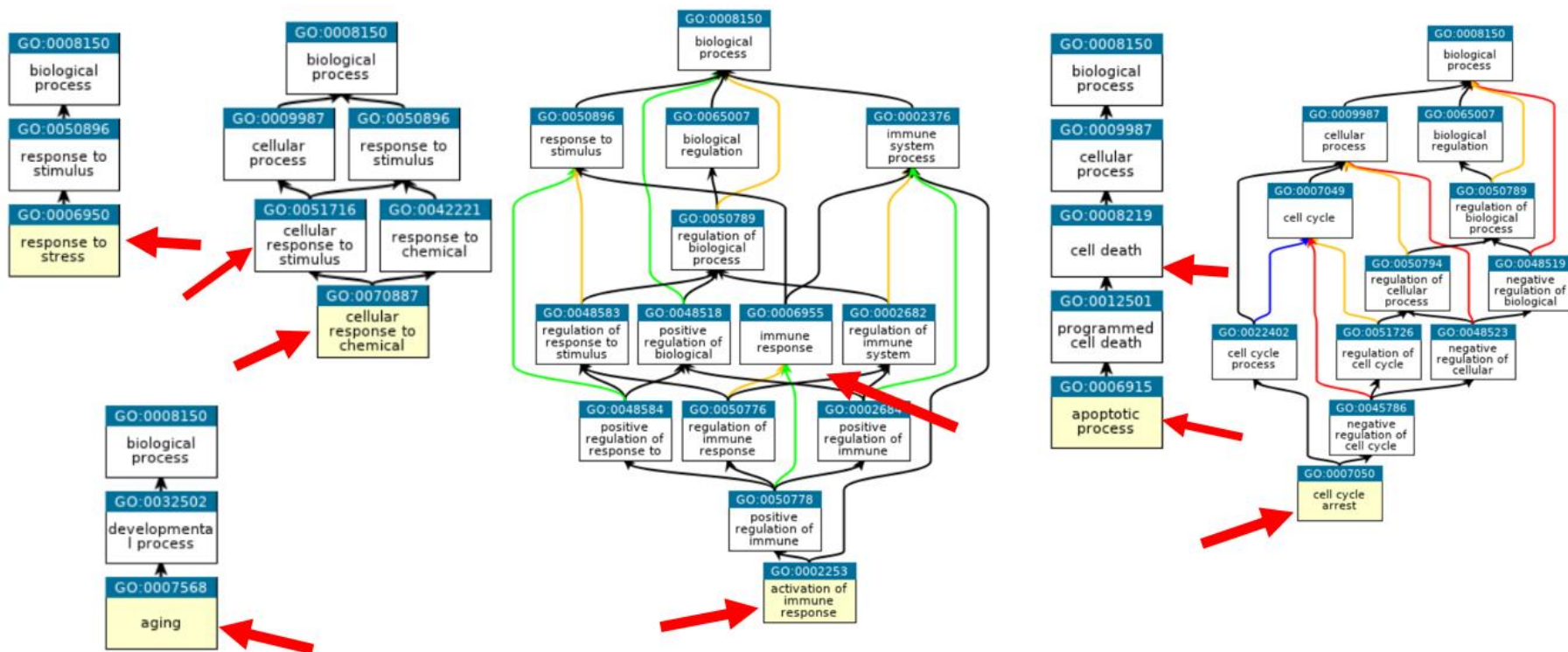
6.1.3 Αποτελέσματα

Τα αποτελέσματα απεικονίζονται στον πίνακα 6.1. Συγκεκριμένα βλέπουμε 24 γονίδια να σχετίζονται με την μικρότερη από τις τιμές του p-value (αυτό πρακτικά σημαίνει ότι είναι σίγουρη η συσχέτιση) με τον μηχανισμό της «κύτταρικής απόκρισης σε χημικό ερέθισμα» ο οποίος αν και δεν υποδεικνύει ξεκάθαρα την ύπαρξη σύνθετων βλαβών στο γενετικό υλικό θα μπορούσε να θεωρηθεί σαν σημαντικό αποτέλεσμα αφού ένα χημικό ερέθισμα θα μπορούσε ενδεχομένως να ειδοποιεί τα κύτταρα για την ύπαρξη σύνθετων βλαβών στο DNA και άρα να συνδέεται με έμμεσο τρόπο με αυτό που μας ενδιαφέρει. Πιο σημαντικοί

μηχανισμοί ως προς την ύπαρξη σύνθετων βλαβών στο DNA είναι ο κυτταρικός θάνατος και η απόπτωση με 19 και 17 γονίδια αντίστοιχα και αρκετά μικρές τιμές p-value ενώ και η «απόκριση σε (κυτταρικό) στρες» αν και πιο γενικός όρος και όχι απαραίτητα κρίσιμος, μπορεί να υποδεικνύει έμμεσα τις σύνθετες βλάβες και εμφανίζει την δεύτερη μικρότερη τιμή p-value και 24 γονίδια να σχετίζονται με αυτόν. Σημαντική είναι και η «ενεργοποίηση της ανοσοαπόκρισης» και ο γενικότερος όρος (βλέπε εικόνα 6.1) «ανοσοαπόκριση» με 11 και 14 γονίδια αντίστοιχα. Επίσης βλέπουμε 27 από τα 29 γονίδια να συμμετέχουν στην «κυτταρική απόκριση σε ερέθισμα» το οποίο αν και γενικός όρος θα μπορούσε να συσχετιστεί με τις σύνθετες βλάβες αφού και αυτές είναι ενός τύπου ερέθισμα για το κύτταρο. Τέλος η γήρανση (του οργανισμού και όχι η κυτταρική γήρανση senescence [53] την οποία εξετάζουμε σε επόμενη ενότητα) και η διακοπή του κυτταρικού κύκλου είναι πολύ σημαντικοί μηχανισμοί που υποδεικνύουν βλάβη στο γενετικό υλικό αλλά εκεί έχουμε μικρό αριθμό γονιδίων και όχι τόσο μικρές τιμές p-value συγκριτικά με τους προηγούμενους μηχανισμούς. Στην εικόνα 6.1 απεικονίζονται με την μορφή δένδρου προγόνων (βλέπε και παράγραφο 4.2) οι βιολογικοί μηχανισμοί που ελέγχθηκαν για να έχει ο αναγνώστης καλύτερη εικόνα της μεταξύ τους σχέσης.

Πίνακας 6.1: Οι βιολογικοί μηχανισμοί που συμμετέχουν τα 29 γονίδια που κρίθηκαν ως σημαντικά και που πιθανόν να σχετίζονται με την ύπαρξη σύνθετων βλαβών στο DNA. Στην πρώτη στήλη έχουμε τους βιολογικούς μηχανισμούς, στην δεύτερη τον αριθμό των γονιδίων που συμμετέχουν στον μηχανισμό, στην τρίτη στήλη το p-value και στην τέταρτη τα ονόματα των γονιδίων.

Term	Count	p-value
cellular response to chemical stimulus	24	1.40E-14
KRAS RAF1 SHC1 BIRC3 CREBBP CASP8 CTNNB1 CDC6 CCNA2 CDK6 ERBB2 FGF1 FGF2 HDAC1 HLA-A MTOR MCM2 MAP3K7 MYD88 PDGFRB STAT1 STAT3 TLR3 TLR4		
response to stress	24	1.90E-11
HRAS KRAS RAF1 SHC1 BIRC3 CASP8 CCNA2 CDK4 CDK6 ERBB2 FGF1 FGF2 HDAC1 HLA-A MTOR MAP3K7 MYD88 NRAS PDGFRB CREBBP STAT1 STAT3 TLR3 TLR4		
cell death	19	6.10E-11
HRAS KRAS RAF1 BIRC3 CASP8 CTNNB1 CDK4 FGF2 HDAC1 MTOR MCM2 MAP3K7 MYD88 PDGFRB STAT1 STAT3 TLR3 TLR4 CREBBP		
apoptotic process	17	1.60E-09
KRAS RAF1 BIRC3 CASP8 CTNNB1 CDK4 HDAC1 MTOR MCM2 MAP3K7 MYD88 PDGFRB STAT1 STAT3 TLR3 TLR4 HRAS		
activation of immune response	11	1.00E-08
HRAS KRAS RAF1 BIRC3 CASP8 MAP3K7 MYD88 NRAS TLR3 TLR4 CREBBP		
cellular response to stimulus	27	1.30E-08
HRAS KRAS RAF1 SHC1 BIRC3 CASP8 CTNNB1 CDC6 CCNA2 CDK4 CDK6 ERBB2 FGF1 FGF2 HDAC1 HLA-A MTOR MCM2 MAP3K7 MYD88 NRAS PDGFRB STAT1 STAT3 TLR3 TLR4 CREBBP		
immune response	14	3.40E-07
HRAS KRAS RAF1 SHC1 BIRC3 CASP8 HLA-A MAP3K7 MYD88 NRAS STAT1 TLR3 TLR4 CREBBP		
aging	6	8.90E-05
KRAS CDK6 FGF2 MTOR STAT3 HRAS		
cell cycle arrest	4	6.80E-03
CDK4 CDK6 MTOR HRAS		



Εικόνα 6.1: Η απεικόνιση των βιολογικών μηχανισμών με την μορφή δένδρου προγόνων. Με κόκκινα βέλη σημειώνονται οι μηχανισμοί που ελέγξαμε για τα 29 γονίδια

6.2 Ασθένειες και τύποι καρκίνου

6.2.1 Στόχος

Θέλουμε για τα 29 γονίδια που κρίναμε ως σημαντικά να εντοπίσουμε τις ασθένειες και τα γενετικά σύνδρομα με τα οποία σχετίζονται. Επίσης ελέγχουμε ξεχωριστά την σχέση των γονιδίων με διάφορους τύπους καρκίνου. Συγκεκριμένα για κάθε γονίδιο σημειώνονται οι τύποι καρκίνου στους οποίους η αυξημένη έκφραση (over-expression) του εκάστοτε γονιδίου έχει στατιστικά σημαντικό ρόλο στην επιβίωση του ασθενούς (hazard-ratio).

6.2.2 Μέθοδος

Στην πλατφόρμα DisGeNet (παράγραφος 4.4.3) ελέγχουμε την συσχέτιση των γονιδίων με ασθένειες. Σαν input δίνουμε την λίστα με τα 29 γονίδια και ως φίλτρο επιλέγουμε το gene disease association (gda) score (δείκτης που παίρνει τιμές από 0 έως 1 και όσο πιο κοντά στην μονάδα βρίσκεται τόσο πιο σίγουρη η συσχέτιση του γονιδίου με την ασθένεια) να είναι τουλάχιστον 0,6. Την ανάλυση για τον καρκίνο πραγματοποιούμε με την βοήθεια της πλατφόρμας GEPIA (παράγραφος 4.4.5) για κάθε γονίδιο ξεχωριστά. Αφού εισάγουμε το όνομα του γονιδίου επιλέγουμε survival analysis και έπειτα ελέγχουμε κάθε έναν από τους τύπους καρκίνου που παρέχει η πλατφόρμα. Για όσους τύπους καρκίνου η υπερέκφραση του γονιδίου δίνει ποσοστό επικινδυνότητας (hazard ratio) μεγαλύτερο της μονάδας και p-value μικρότερο του 0,01 το σημειώνουμε ως στατιστικά σημαντικό για την ανάλυση μας. Δηλαδή η υπερέκφραση του συγκεκριμένου γονιδίου σε ασθενείς με τον εκάστοτε τύπο καρκίνου σχετίζεται με χαμηλά ποσοστά επιβίωσης του ασθενούς και θα μπορούσε θεωρητικά να αποτελέσει ένδειξη του κινδύνου που διατρέχει ο ασθενής ή ακόμα και διαγνωστικό δείκτη για τον συγκεκριμένο τύπο καρκίνου.

6.2.3 Αποτελέσματα

Στον πίνακα 6.2 συγκεντρώνονται οι ασθένειες με τις οποίες σχετίζονται τα 29 γονίδια. Στον πίνακα φαίνεται και το ID της κάθε ασθένειας, το gene - disease association (gda) score καθώς και το πότε ήταν η πρώτη και η τελευταία φορά που αναφέρθηκε η συσχέτιση σε κάποια δημοσίευση. Από τα 29 γονίδια τα 19 εμφανίζουν έστω μια συσχέτιση με κάποια ασθένεια ή σύνδρομο η οποία είναι πιθανό να ισχύει αφού το gda score είναι τουλάχιστον 0,6. Αυτό που παρατηρείται στον πίνακα είναι οι πολλές μορφές καρκίνου με τις οποίες σχετίζονται τα γονίδια, κάτι που υποδεικνύει την συσχέτιση τους με την ύπαρξη σύνθετων βλαβών στο DNA αφού ένα συχνό επακόλουθο των βλαβών αυτών είναι η δημιουργία καρκινικών όγκων.

Πίνακας 6.2: Ασθένειες και σύνδρομα με τα οποία σχετίζονται τα 29 γονίδια που προκρίθηκαν ως σημαντικά. Στις στήλες του πίνακα έχουμε με την σειρά: το όνομα του γονιδίου, το όνομα της ασθένειας, το Disease ID κάθε ασθένειας, το gene disease association score, την πρώτη και την τελευταία αναφορά της συσχέτισης.

Gene	Disease	Disease_id	Score_gda	First Report	Last Report
BIRC3	Mucosa-Associated Lymphoid Tissue Lymphoma	C0242647	0.6	1999	2017
CASP8	Autoimmune Lymphoproliferative Syndrome Type 2B	C1846545	0.92	2002	2009
CASP8	Breast Carcinoma	C0678222	0.7	2004	2018
CASP8	Liver carcinoma	C2239176	0.7	1999	2017
CASP8	Malignant neoplasm of breast	C0006142	0.6	2004	2016
CDK4	Melanoma	C0025202	0.7	1994	2018
CDK4	Liposarcoma	C0023827	0.6	1995	2016
CDK4	Liposarcoma, Dedifferentiated	C0205824	0.6	2003	2017
CDK6	Primary autosomal recessive microcephaly 12	C4015156	0.6	2014	2014
CREBBP	Rubinstein-Taybi Syndrome	C0035934	1	1995	2018
CREBBP	Leukemia, Myelocytic, Acute	C0023467	0.6	1996	2018

CTNNB1	Ovarian neoplasm	C0919267	0.84	1999	2016
CTNNB1	Malignant neoplasm of ovary	C1140680	0.81	1999	2014
CTNNB1	Medulloblastoma	C0025149	0.8	1998	2016
CTNNB1	Liver carcinoma	C2239176	0.8	1998	2017
CTNNB1	Pilomatrixoma	C0206711	0.75	1997	2015
CTNNB1	Fibromatosis, Aggressive	C0079218	0.7	1997	2017
CTNNB1	Hepatoblastoma	C0206624	0.7	1999	2016
CTNNB1	Mental retardation, autosomal dominant 19	C3554449	0.7	2012	2017
CTNNB1	Craniopharyngioma	C0010276	0.66	2004	2016
CTNNB1	Malignant neoplasm of prostate	C0376358	0.6	2000	2018
CTNNB1	Colorectal Cancer	C1527249	0.6	1997	2017
ERBB2	Malignant neoplasm of breast	C0006142	0.8	1987	2018
ERBB2	Non-Small Cell Lung Carcinoma	C0007131	0.7	1989	2018
ERBB2	Stomach Neoplasms	C0038356	0.7	1989	2015
ERBB2	Adenocarcinoma of lung	C0152013	0.7	1991	2017
ERBB2	Stomach Carcinoma	C0699791	0.7	1986	2018
ERBB2	Ovarian neoplasm	C0919267	0.7	1990	2017
ERBB2	Malignant neoplasm of ovary	C1140680	0.7	1989	2018
ERBB2	Mammary Neoplasms	C1458155	0.7	1987	2018
ERBB2	Malignant neoplasm of stomach	C0024623	0.6	1986	2018
ERBB2	Breast Carcinoma	C0678222	0.6	1987	2018
HRAS	Costello syndrome	C0587248	1	1982	2018
HRAS	Nevus sebaceous	C3854181	0.72	2012	2018
HRAS	Malignant neoplasm of urinary bladder	C0005684	0.7	1983	2015
HRAS	Melanoma	C0025202	0.7	1984	2015
HRAS	Carcinoma of bladder	C0699885	0.7	1983	2016
HRAS	Bladder Neoplasm	C0005695	0.6	1983	2015
HRAS	Malignant neoplasm of breast	C0006142	0.6	1986	2011
HRAS	Squamous cell carcinoma	C0007137	0.6	1986	2016
HRAS	Neuroblastoma	C0027819	0.6	1988	2015
HRAS	Breast Carcinoma	C0678222	0.6	1986	2011
HRAS	Myopathy, congenital, with excess of muscle spindles	C1968782	0.6	2005	2017
KRAS	Juvenile Myelomonocytic Leukemia	C0349639	1	1987	2017

KRAS	Noonan Syndrome 3	C1860991	0.9	2006	2015
KRAS	Pilocytic Astrocytoma	C0334583	0.72	2005	2013
KRAS	Non-Small Cell Lung Carcinoma	C0007131	0.7	1989	2018
KRAS	Leukemia, Myelocytic, Acute	C0023467	0.7	1988	2017
KRAS	Noonan Syndrome	C0028326	0.7	2002	2016
KRAS	Pancreatic Neoplasm	C0030297	0.7	1993	2016
KRAS	Adenocarcinoma of lung	C0152013	0.7	1990	2018
KRAS	Pancreatic carcinoma	C0235974	0.7	1986	2018
KRAS	Malignant neoplasm of lung	C0242379	0.7	1990	2018
KRAS	Malignant neoplasm of pancreas	C0346647	0.7	1992	2018
KRAS	Breast Carcinoma	C0678222	0.7	1988	2016
KRAS	Carcinoma of lung	C0684249	0.7	1983	2018
KRAS	Ovarian neoplasm	C0919267	0.7	1991	2013
KRAS	Cardio-facio-cutaneous syndrome	C1275081	0.7	2006	2017
KRAS	Hemangiosarcoma	C0018923	0.64	1994	2016
KRAS	Thyroid Neoplasm	C0040136	0.63	1990	2016
KRAS	Malignant neoplasm of urinary bladder	C0005684	0.62	1984	2018
KRAS	Nevus sebaceous	C3854181	0.62	2012	2017
KRAS	Malignant neoplasm of breast	C0006142	0.6	1988	2016
KRAS	Malignant Neoplasms	C0006826	0.6	2000	2018
KRAS	Lung Neoplasms	C0024121	0.6	1993	2017
KRAS	Malignant neoplasm of ovary	C1140680	0.6	1997	2017
KRAS	Colorectal Cancer	C1527249	0.6	1993	2018
MAP3K7	Forney Robinson Pascoe syndrome	C2931461	0.7	2010	2016
MCM2	Deafness, autosomal dominant 70	C4310775	0.6	2016	2016
MTOR	Focal cortical dysplasia of Taylor	C1846385	0.62	2015	2017
MTOR	Conventional (Clear Cell) Renal Cell Carcinoma	C0279702	0.6	2009	2018
MTOR	Smith-Kingsmore syndrome	C4225259	0.6	2016	2016
MYD88	Waldenstrom Macroglobulinemia	C0024419	0.8	2012	2018
MYD88	MYD88 Deficiency	C2677092	0.71	2008	2014
MYD88	Diffuse Large B-Cell Lymphoma	C0079744	0.6	2013	2018

NRAS	Juvenile Myelomonocytic Leukemia	C0349639	0.8	1985	2017
NRAS	Noonan Syndrome	C0028326	0.77	2004	2015
NRAS	Leukemia, Myelocytic, Acute	C0023467	0.7	1985	2017
NRAS	Melanoma	C0025202	0.7	1988	2018
NRAS	Congenital giant melanocytic nevus	C1842036	0.7	1987	2014
NRAS	Nevus sebaceous	C3854181	0.7	2012	2012
NRAS	Epidermal nevus	C0334082	0.6	2003	2014
NRAS	Colorectal Cancer	C1527249	0.6	1993	2018
NRAS	RAS-associated autoimmune leukoproliferative disorder	C2674723	0.6	1985	2011
NRAS	Noonan Syndrome 6	C2750732	0.6	1985	2015
PDGFRB	Infantile myofibromatosis	C0432284	0.94	2004	2017
PDGFRB	Fahr's syndrome	C0393590	0.71	2015	2015
PDGFRB	Kosaki overgrowth syndrome	C4225270	0.7	2015	2015
PDGFRB	Penttinen-Aula syndrome	C1866182	0.6	2015	2015
PDGFRB	Basal ganglia calcification, idiopathic, 4	C3554321	0.6	1993	2016
RAF1	Noonan Syndrome 5	C1969057	0.81	2007	2016
RAF1	Noonan Syndrome	C0028326	0.8	1991	2016
RAF1	Hypertrophic Cardiomyopathy	C0007194	0.7	2007	2017
RAF1	Noonan syndrome with multiple lentigines	C0175704	0.64	1993	2012
RAF1	LEOPARD syndrome 2	C1969056	0.6	2007	2015
RAF1	Cardiomyopathy, dilated, 1NN	C4014656	0.6	2014	2015
STAT1	Mycobacterial and viral infections, susceptibility to, autosomal recessive	C3151088	0.7	2003	2017
STAT1	Immunodeficiency 31C	C3279990	0.7	2011	2016
STAT1	Immunodeficiency 31A	C4013950	0.7	2001	2012
STAT1	Malignant neoplasm of breast	C0006142	0.6	2000	2018
STAT3	Hyper-Immunoglobulin E	C2936739	0.9	2007	2016

	Syndrome, Autosomal Dominant				
STAT3	Job Syndrome	C3887645	0.8	2004	2016
STAT3	Hyper-Ige Recurrent Infection Syndrome, Autosomal Dominant	C3489795	0.7	1966	2018
STAT3	Autoimmune disease, multisystem, infantile-onset, 1	C4014795	0.7	2015	2017
STAT3	Psoriasis	C0033860	0.69	2005	2017
TLR4	Necrotizing Enterocolitis	C0520459	0.6	2006	2015

Στον πίνακα 6.3 έχουμε συγκεντρώσει τα αποτελέσματα της ανάλυσης για τύπους καρκίνου από την πλατφόρμα GEPIA. Κάθε στήλη του πίνακα αντιστοιχεί σε ένα από τα 29 γονίδια ενώ κάθε γραμμή αντιστοιχεί σε κάποιον τύπο καρκίνου. Συγκεκριμένα: ACC: Adrenocortical carcinoma, BLCA: Bladder Urothelial Carcinoma, BRCA: Breast invasive carcinoma, CESC: Cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL: Cholangio carcinoma, COAD: Colon adenocarcinoma, DLBC: Lymphoid Neoplasm Diffuse Large B-cell Lymphoma, ESCA: Esophageal carcinoma, GBM: Glioblastoma multiforme, HNSC: Head and Neck squamous cell carcinoma, KICH: Kidney Chromophobe, KIRC: Kidney renal clear cell carcinoma, KIRP: Kidney renal papillary cell carcinoma, LAML: Acute Myeloid Leukemia, LGG: Brain Lower Grade Glioma, LIHC: Liver hepatocellular carcinoma, LUAD: Lung adenocarcinoma, LUSC: Lung squamous cell carcinoma, MESO: Mesothelioma, OV: Ovarian serous cystadenocarcinoma, PAAD: Pancreatic adenocarcinoma, PCPG: Pheochromocytoma and Paraganglioma, PRAD: Prostate adenocarcinoma, READ: Rectum adenocarcinoma, SARC: Sarcoma, SKCM: Skin Cutaneous Melanoma, STAD: Stomach adenocarcinoma, TGCT: Testicular Germ Cell

Tumors, THCA: Thyroid carcinoma, THYM: Thymoma, UCEC: Uterine Corpus Endometrial Carcinoma, UCS: Uterine Carcinosarcoma, UVM: Uveal Melanoma

Για όσους τύπους καρκίνου προκύπτει ποσοστό επικινδυνότητας μεγαλύτερο της μονάδας και p-value μικρότερο από 0,01 σημειώνεται με κουκίδα το αντίστοιχο κελί του πίνακα. Βλέπουμε ότι για κάποιους τύπους καρκίνου, όπως ο καρκίνος του μαστού (BRCA) η υπερέκφραση κανενός από τα 29 γονίδια δεν έχει στατιστικά σημαντική συσχέτιση με την επιβίωση του ασθενούς ενώ για άλλους όπως το γλοίωμα του εγκεφάλου (LGG) τα περισσότερα γονίδια παρουσιάζουν στατιστικά σημαντική συσχέτιση ανάμεσα στην υπερέκφραση τους και την επιβίωση του ασθενούς.

	CASP8	CDC20	CCNA2	SHC1	MCM2	MCM4	HDAC1	CDK4	CDK6	KRAS	RAF1	CREBBP	HRAS	CDC6	NRAS	PDGFRB	MAP3K7	BIRC3	MYD88	ERBB2	FGF2	FGF1	TLR4	MTOR	HLA-A	STAT3	CTNNB1	STAT1	TLR3	
ACC		•	•			•			•				•	•										•						
BLCA																						•	•							
BRCA																														
CESC				•					•																					
CHOL																														
COAD																														
DLBC																														
ESCA																														
GBM																			•											
HNSC																														
KICH																														
KIRC		•	•																											
KIRP		•	•																											
LAML					•												•													
LGG	•	•	•	•	•	•	•		•					•	•	•			•	•		•			•	•		•	•	
LIHC		•	•		•			•					•														•	•		•
LUAD		•	•	•		•								•	•															
LUSC																	•													
MESO		•	•	•	•	•		•	•	•				•																
OV																														
PAAD			•						•					•																
PCPG																														
PRAD																														
READ																														
SARC			•					•	•						•		•							•						
SKCM		•				•		•					•	•																
STAD																														
TGCT																								•						
THCA																														
THYM																														
UCEC																														
UCS																														
UVM				•													•									•				

Πίνακας 6.3: Ο πίνακας της ανάλυσης των 29 γονιδίων στην πλατφόρμα GEPIA. Κάθε στήλη του πίνακα είναι ένα από τα 29 γονίδια και κάθε γραμμή ένας τύπος καρκίνου. Όταν η υπερέκφραση κάποιου από τα γονίδια σχετίζεται με στατιστικώς σημαντικά (p-value < 0,01) ποσοστά χαμηλής επιβίωσης του ασθενούς με τον συγκεκριμένο τύπο καρκίνου, σημειώνουμε με κουκίδα το αντίστοιχο κελί.

6.3 Κυτταρική γήρανση (*senescence*)

6.3.1 Στόχος

Θέλουμε να ελέγξουμε αν και για ποια από τα 29 γονίδια που κρίθηκαν ως σημαντικά υπάρχει συσχέτιση με τον βιολογικό μηχανισμό της κυτταρικής γήρανσης. Η κυτταρική γήρανση ή *senescence* είναι μια κατάσταση αδράνειας με ταυτόχρονη έκκριση μορίων από το εσωτερικό του κυττάρου στην οποία ένα κύτταρο μπορεί να περιέλθει έπειτα από κάποιο ζημιογόνο για αυτό γεγονός. Η πρόκληση σύνθετων βλαβών στο γενετικό υλικό των κυττάρων μετά από ακτινοβολία είναι ένα στρεσογόνο γεγονός το οποίο είναι πιθανό να οδηγήσει το κύτταρο στην κατάσταση της γήρανσης. Επομένως αντιλαμβάνεται ο αναγνώστης ότι αν υπάρχει πράγματι συσχέτιση κάποιων από τα 29 γονίδια με την κυτταρική γήρανση είναι πιθανό τα ίδια γονίδια να ενεργοποιούνται με την ύπαρξη σύνθετων βλαβών στο DNA.

6.3.2 Μέθοδος

Χρησιμοποιούμε την πλατφόρμα βιοπληροφορικής ανάλυσης του καθηγητή Βασιλή Γοργούλη SeneQuest. Η πλατφόρμα αυτή (βλέπε παράγραφο 4.4.6) δημιουργήθηκε ακριβώς για αυτόν τον σκοπό: να συγκεντρώσει όση πληροφορία υπάρχει διαθέσιμη στην βιβλιογραφία για την συσχέτιση γονιδίων με την κατάσταση *senescence*. Η λειτουργία του είναι εξαιρετικά απλή και εύχρηστη: λαμβάνεται σαν input το όνομα καθενός από τα 29 γονίδια και το output είναι ο αριθμός των δημοσιεύσεων στις οποίες αναφέρεται η συσχέτιση του με το term *senescence* καθώς και το αν η έκφραση του αυξάνεται ή μειώνεται στην συγκεκριμένη κατάσταση σύμφωνα με την εκάστοτε δημοσίευση. Κρατάμε προφανώς μόνο όσα αποτελέσματα έχουν να κάνουν με τον άνθρωπο και όχι άλλους οργανισμούς.

6.3.3 Αποτελέσματα

Συγκεντρώνουμε τα αποτελέσματα στον πίνακα 6.4. Έχουμε το όνομα του γονιδίου και δίπλα τον αριθμό των δημοσιεύσεων που αναφέρουν συσχέτιση είτε με υπερέκφραση (up-regulated), είτε με υποέκφραση (down-regulated) του γονιδίου κατά την κατάσταση της κυτταρικής γήρανσης. Στην περίπτωση που για κάποιο γονίδιο δεν υπάρχει καμία αναφορά συσχέτισης με την κατάσταση της γήρανσης βάζουμε παύλα. Το ότι για κάποια γονίδια υπάρχει αναφορά τόσο για αύξηση όσο και για μείωση της έκφρασης τους, έχει να κάνει με τον τύπο του κυττάρου που εξετάζεται κάθε φορά. Βλέπουμε ότι για 26 από τα 29 γονίδια υπάρχει τουλάχιστον μια αναφορά σε δημοσιεύσεις που επιβεβαιώνει την συσχέτιση τους με την κατάσταση senescence γεγονός που υποδεικνύει την πιθανή συσχέτιση τους με την ύπαρξη σύνθετων βλαβών στο γενετικό υλικό.

Πίνακας 6.4: Ο αριθμός των δημοσιεύσεων που συσχετίζει κάθε ένα από τα 29 γονίδια με την κατάσταση της κυτταρικής γήρανσης. Ανάλογα το αν σε αυτή την κατάσταση αναφέρεται αύξηση ή μείωση της έκφρασης του εκάστοτε γονιδίου συμπληρώνεται η στήλη up-regulated ή down-regulated αντίστοιχα.

Gene	Up-regulated reports	Down-regulated reports
CASP8		1
CDC20		5
CCNA2	1	5
SHC1	4	
MCM2		1
MCM4	~	~
HDAC1	2	2
CDK4		7
CDK6		3
KRAS	5	2
RAF1	2	
CREBBP	1	1
HRAS	25	
CDC6	3	
NRAS	4	
PDGFRB		1
MAP3K7	2	
BIRC3	1	1
MYD88	~	~
ERBB2	5	
FGF2	2	3
FGF1	2	
TLR4	2	
MTOR	23	2
HLA-A	~	~
STAT3	11	2
CTNNB1	3	2
STAT1	5	1
TLR3	3	

6.4 Συνοψιση αποτελεσμάτων - συμπεράσματα

Στο προηγούμενο κεφάλαιο επιλέχθηκαν 29 γονίδια ως τα πλέον σημαντικά για περαιτέρω ανάλυση. Ξεκινώντας λοιπόν την ανάλυση σε αυτό το κεφάλαιο στόχος ήταν να εξακριβωθεί κατά πόσο τα συγκεκριμένα γονίδια συσχετίζονται πραγματικά με την ύπαρξη σύνθετων βλαβών DNA και άρα κατά πόσο ορθώς επιλέχθηκαν ως σημαντικά. Από τα αποτελέσματα της επεξεργασίας με τα διαφορετικά εργαλεία βιοπληροφορικής γίνεται φανερό ότι υπάρχει πολύ μεγάλη πιθανότητα αυτά τα γονίδια να ενεργοποιούνται πράγματι όταν στο κύτταρο εντοπίζονται σύνθετες βλάβες DNA. Αρχικά στην ανάλυση για τους βιολογικούς μηχανισμούς στους οποίους συμμετέχουν τα 29 γονίδια, εντοπίστηκαν εννέα διαφορετικοί μηχανισμοί με αρκετά χαμηλές τιμές p-value που «μαρτυρούν» είτε άμεσα είτε με πιο έμμεσο τρόπο την ύπαρξη σύνθετων βλαβών στο γενετικό υλικό. Στην συνέχεια αναζητήθηκαν οι ασθένειες που σχετίζονται με τα συγκεκριμένα γονίδια (προέκυψαν αποτελέσματα για 19 από τα 29) είτε μέσω κάποιας μετάλλαξης, είτε μέσω αλλαγής στην έκφραση του γονιδίου και το γεγονός ότι η πλειοψηφία των ασθενειών ήταν νεοπλασίες και τύποι καρκίνου υποδεικνύουν την ύπαρξη σύνθετων βλαβών στο γενετικό υλικό, αφού η αδυναμία του κυττάρου να επιδιορθώσει το πρόβλημα οδηγεί συχνά σε ανεξέλεγκτο πολλαπλασιασμό και δημιουργία καρκινικών όγκων. Παρατηρήθηκε επίσης ότι για 26 από τα 29 γονίδια, η αυξημένη έκφραση τους έχει στατιστικά σημαντική συσχέτιση με την επιβίωση του ασθενούς σε διαφορετικούς τύπους καρκίνου ενώ επίσης για 26 από τα 29 (διαφορετικά από πριν) υπάρχουν αναφορές για συσχέτιση τους με την κυτταρική γήρανση (senescence), κατάσταση που επίσης είναι πιθανό να προκληθεί στο κύτταρο όταν υπάρχουν σύνθετες βλάβες στο DNA. Το γεγονός ότι και τα τέσσερα εργαλεία βιοπληροφορικής ανάλυσης έδωσαν αποτελέσματα που υποδεικνύουν την συσχέτιση των 29 γονιδίων με τις σύνθετες βλάβες DNA αυξάνει σημαντικά τις πιθανότητες να ισχύει πράγματι η συσχέτιση και επομένως επιβεβαιώνει την επιλογή να αναλυθούν τα συγκεκριμένα γονίδια. Σημαντική εξαίρεση αποτελεί το γονίδιο MCM4 το οποίο δεν εμφανίζεται να συμμετέχει σε κανέναν από τους 9 βιολογικούς μηχανισμούς που εντοπίστηκαν από την

ανάλυση στην πλατφόρμα DAVID, δεν υπάρχει αναφορά για συσχέτιση του με την κυτταρική γήρανση όπως επίσης δεν υπάρχει αναφορά για συσχέτιση του με κάποια ασθένεια μέσω της πλατφόρμας DisGeNet. Μόνο η ανάλυση στην GEPiA έδωσε κάποια αποτελέσματα για το συγκεκριμένο γονίδιο, το οποίο συσχετίζεται με 5 διαφορετικούς τύπους καρκίνου, όμως το γεγονός ότι οι τρεις από τις τέσσερις πλατφόρμες δεν έδωσαν κάποιο αποτέλεσμα υποδεικνύει ότι το MCM4 ενδεχομένως να μην σχετίζεται με τις σύνθετες βλάβες DNA.

Κεφάλαιο 7: Συμπεράσματα της διπλωματικής εργασίας και μελλοντικές κατευθύνσεις

Αντικείμενο της παρούσας μελέτης είναι ο μη καθιερωμένος μέχρι στιγμής όρος γονιδιακής οντολογίας «Απόκριση στις σύνθετες βλάβες DNA»

Μετά από την κατάλληλη αναζήτηση και φιλτράρισμα των αποτελεσμάτων προέκυψαν 29 γονίδια για τα οποία οι πιθανότητες να ενεργοποιούνται όντως με την εμφάνιση σύνθετων βλαβών στο DNA είναι μεγάλες.

Ακολούθησε περαιτέρω επεξεργασία αυτών των γονιδίων με ειδικές πλατφόρμες βιοπληροφορικής ανάλυσης. Χρησιμοποιήθηκαν τέσσερις διαφορετικές πλατφόρμες και στα αποτελέσματα που έδινε η κάθε μια από αυτές αναζητήθηκαν ενδείξεις για την συσχέτιση των γονιδίων με τις σύνθετες βλάβες DNA. Τα αποτελέσματα επιβεβαίωσαν την συσχέτιση αυτή.

- Εντοπίσαμε εννέα βιολογικούς μηχανισμούς που να υποδεικνύουν την ύπαρξη σύνθετων βλαβών DNA στο κύτταρο. Οι μηχανισμοί αυτοί είναι: «κυτταρική απόκριση σε χημικό ερέθισμα», «απόκριση σε στρες», «κυτταρικός θάνατος», «αποπτωτική διαδικασία», «ενεργοποίηση της ανοσοαπόκρισης», «κυτταρική απόκριση σε ερέθισμα», «ανοσοαπόκριση», «γήρανση», «διακοπή κυτταρικού κύκλου».
- Για 19 από τα 29 γονίδια υπάρχει συσχέτιση με διάφορες ασθένειες οι περισσότερες από τις οποίες πρόκειται για νεοπλασίες.

- 26 από τα 29 γονίδια επηρεάζουν την επιβίωση του ασθενούς σε διαφορετικούς τύπους καρκίνου.
- Για 26 από τα 29 υπάρχουν αναφορές για συσχέτιση τους με την κατάσταση της κυτταρικής γήρανσης.
- Παρατηρήσαμε ότι 15 από τα 29 γονίδια εμφανίζονται στα αποτελέσματα και των τεσσάρων αναλύσεων στις πλατφόρμες βιοπληροφορικής επεξεργασίας. Αυτά είναι τα CASP8, MCM2, CDK4, CDK6, KRAS, HRAS, NRAS, CREBBP, PDGFRB, MAP3K7, BIRC3, ERBB2, MTOR, STAT1 και STAT3 και μπορούμε να πούμε ότι αυτά έχουν την μεγαλύτερη πιθανότητα για ενδεχόμενη συσχέτιση με τις σύνθετες βλάβες DNA.

Μελλοντικές κατευθύνσεις

Προτείνουμε, βασιζόμενοι στην προηγούμενη ανάλυση, περαιτέρω διερεύνηση των 15 γονιδίων που εντοπίσαμε ως τα πλέον πιθανά υποψήφια για συμμετοχή στην απόκριση των ανθρώπινων κυττάρων στις σύνθετες βλάβες DNA. Για τα 15 γονίδια που εμφανίζονται και στις 4 αναλύσεις που πραγματοποιήθηκαν είναι σημαντικό να μελετηθεί η λειτουργία τους και να προσδιοριστεί το αν συμμετέχουν σε κάποιον ή κάποιους από τους γνωστούς επιδιορθωτικούς μηχανισμούς του DNA.

Για το γονίδιο MCM4 θα ήταν χρήσιμο να γίνουν μελέτες ώστε να αποκαλυφθεί η ακριβής λειτουργία του και να εξακριβωθεί το αν τελικά και με ποιον τρόπο συμμετέχει στην απόκριση του κυττάρου στις σύνθετες βλάβες DNA.

Επίσης πολύ χρήσιμη θα ήταν μια διαφορετική προσέγγιση (με διαφορετικές μεθόδους από αυτές που χρησιμοποιήθηκαν στην παρούσα εργασία) στο φιλτράρισμα των 869 γονιδίων από τα οποία ξεκινήσαμε αρχικά αλλά και χρήση διαφορετικών εργαλείων βιοπληροφορικής ανάλυσης ώστε να προκύψουν, αν φυσικά υπάρχουν και άλλα γονίδια που συμμετέχουν στην απόκριση των κυττάρων στις σύνθετες βλάβες DNA. Τέλος αφού συγκεντρωθούν όλα αυτά τα

γονίδια (και με την προσθήκη προφανώς και όλων εκείνων που είναι ήδη γνωστά ότι ανήκουν σε κάποιον μηχανισμό επιδιόρθωσης) θα μπορεί να δημιουργηθεί ένας νέος όρος γονιδιακής οντολογίας με όνομα «απόκριση στις **σύνθετες** βλάβες DNA» που θα είναι χρήσιμος σε μελλοντικές έρευνες που θα επικεντρώνονται στον συγκεκριμένο τομέα.

Βιβλιογραφία

1. Bianconi, E., et al., *An estimation of the number of cells in the human body*. Ann Hum Biol, 2013. **40**(6): p. 463-71.
2. Wolfe, S.L., *Biology of the cell*. 1972, Belmont, Calif.: Wadsworth. 545 p.
3. NCBI. *What is a cell? A Basic Introduction to the Science Underlying NCBI Resources* 2004; Available from: https://web.archive.org/web/20130507094245/http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html.
4. Alberts, B., *Essential cell biology*. Fourth edition. ed. 2014, New York, NY: Garland Science. 1 volume (various pagings).
5. Slupphaug, G., B. Kavli, and H.E. Krokan, *The interacting pathways for prevention and repair of oxidative DNA damage*. Mutat Res, 2003. **531**(1-2): p. 231-51.
6. De Bont, R. and N. van Larebeke, *Endogenous DNA damage in humans: a review of quantitative data*. Mutagenesis, 2004. **19**(3): p. 169-85.
7. Nikitaki, Z., et al., *Measurement of complex DNA damage induction and repair in human cellular systems after exposure to ionizing radiations of varying linear energy transfer (LET)*. Free Radic Res, 2016. **50**(sup1): p. S64-S78.
8. Georgakilas, A.G., P. O'Neill, and R.D. Stewart, *Induction and repair of clustered DNA lesions: what do we know so far?* Radiat Res, 2013. **180**(1): p. 100-9.
9. Mazouzi, A., G. Velimezi, and J.I. Loizou, *DNA replication stress: causes, resolution and disease*. Exp Cell Res, 2014. **329**(1): p. 85-93.
10. World Health Organisation, *Ionizing radiation, health effects and protective measures*. 2016; Available from: <https://www.who.int/news-room/fact-sheets/detail/ionizing-radiation-health-effects-and-protective-measures>.
11. Eccles, L.J., P. O'Neill, and M.E. Lomax, *Delayed repair of radiation induced clustered DNA damage: friend or foe?* Mutat Res, 2011. **711**(1-2): p. 134-41.
12. IAEA, *RADIATION BIOLOGY: A HANDBOOK FOR TEACHERS AND STUDENTS*. 2010, Vienna: IAEA.
13. Μαργαρίτης Λ., *Ραδιοβιολογία - Ακτινοβολίες Και Ζωή*. 1996: Συμμετρία.
14. Elgazzar, A.H. and N. Kazem, *Biological Effects of Ionizing Radiation, in The Pathophysiological Basis of Nuclear Medicine*. 2006. p. 540-548.
15. Bennett, P.V., et al., *Are endogenous clustered DNA damages induced in human cells?* Free Radic Biol Med, 2004. **37**(4): p. 488-99.
16. Terzidis, M.A., C. Ferreri, and C. Chatgililoglu, *Radiation-induced formation of purine lesions in single and double stranded DNA: revised quantification*. Front Chem, 2015. **3**: p. 18.
17. Ward, J.F., *Some Biochemical Consequences of the Spatial Distribution of Ionizing Radiation-Produced Free Radicals*. Radiation Research, 1981. **86**(2).
18. Sage, E. and N. Shikazono, *Radiation-induced clustered DNA lesions: Repair and mutagenesis*. Free Radic Biol Med, 2017. **107**: p. 125-135.
19. Watson JD, B.T., Bell SP, Gann A, Levine M, Losick R, *Molecular Biology of the Gene (5th ed.)*. 2004: Pearson Benjamin Cummings; CSHL Press.

20. Giglia-Mari, G., A. Zotter, and W. Vermeulen, *DNA damage response*. Cold Spring Harb Perspect Biol, 2011. **3**(1): p. a000745.
21. Almeida, K.H. and R.W. Sobol, *A unified view of base excision repair: lesion-dependent protein complexes regulated by post-translational modification*. DNA Repair (Amst), 2007. **6**(6): p. 695-711.
22. Scharer, O.D., *Nucleotide excision repair in eukaryotes*. Cold Spring Harb Perspect Biol, 2013. **5**(10): p. a012609.
23. Marteijn, J.A., et al., *Understanding nucleotide excision repair and its roles in cancer and ageing*. Nat Rev Mol Cell Biol, 2014. **15**(7): p. 465-81.
24. Muro, Y., et al., *DNA mismatch repair enzymes: genetic defects and autoimmunity*. Clin Chim Acta, 2015. **442**: p. 102-9.
25. Li, G.M., *Mechanisms and functions of DNA mismatch repair*. Cell Res, 2008. **18**(1): p. 85-98.
26. Sung, P. and H. Klein, *Mechanism of homologous recombination: mediators and helicases take on regulatory functions*. Nat Rev Mol Cell Biol, 2006. **7**(10): p. 739-50.
27. Deriano, L. and D.B. Roth, *Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage*. Annu Rev Genet, 2013. **47**: p. 433-55.
28. Mavragani, I.V., et al., *Complex DNA Damage: A Route to Radiation-Induced Genomic Instability and Carcinogenesis*. Cancers (Basel), 2017. **9**(7).
29. Lomax, M.E., L.K. Folkes, and P. O'Neill, *Biological consequences of radiation-induced DNA damage: relevance to radiotherapy*. Clin Oncol (R Coll Radiol), 2013. **25**(10): p. 578-85.
30. Sutherland, B.M., et al., *Clustered Damages and Total Lesions Induced in DNA by Ionizing Radiation: Oxidized Bases and Strand Breaks*. Biochemistry, 2000. **39**(27): p. 8026-8031.
31. Nikitaki, Z., et al., *Non-DSB clustered DNA lesions. Does theory colocalize with the experiment?* Radiation Physics and Chemistry, 2016. **128**: p. 26-35.
32. Kuo, L.J. and L.X. Yang, *Gamma-H2AX - a novel biomarker for DNA double-strand breaks*. In Vivo, 2008. **22**(3): p. 305-9.
33. Nakad, R. and B. Schumacher, *DNA Damage Response and Immune Defense: Links and Mechanisms*. Front Genet, 2016. **7**: p. 147.
34. Rodier, F., et al., *Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion*. Nat Cell Biol, 2009. **11**(8): p. 973-9.
35. Xu, Y., *DNA damage: a trigger of innate immunity but a requirement for adaptive immune homeostasis*. Nat Rev Immunol, 2006. **6**(4): p. 261-70.
36. Luscombe, N.M., D. Greenbaum, and M. Gerstein, *What is bioinformatics? A proposed definition and overview of the field*. Methods Inf Med, 2001. **40**(4): p. 346-58.
37. Xiong, J., *Essential bioinformatics*. 2006, Cambridge ; New York: Cambridge University Press. xi, 339 p.
38. Sim, A.Y., P. Minary, and M. Levitt, *Modeling nucleic acids*. Curr Opin Struct Biol, 2012. **22**(3): p. 273-8.
39. Gene Ontology, C., *The Gene Ontology project in 2008*. Nucleic Acids Res, 2008. **36**(Database issue): p. D440-4.

40. Koonin, E.V., *Orthologs, paralogs, and evolutionary genomics*. *Annu Rev Genet*, 2005. **39**: p. 309-38.
41. NCBI Resource Coordinators, *Database resources of the National Center for Biotechnology Information*. *Nucleic Acids Res*, 2018. **46**(D1): p. D8-D13.
42. Carbon, S., et al., *AmiGO: online access to ontology and annotation data*. *Bioinformatics*, 2009. **25**(2): p. 288-9.
43. Yates, A.D., et al., *Ensembl 2020*. *Nucleic Acids Res*, 2020. **48**(D1): p. D682-D688.
44. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. *Nat Protoc*, 2009. **4**(1): p. 44-57.
45. Szklarczyk, D., et al., *STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets*. *Nucleic Acids Res*, 2019. **47**(D1): p. D607-D613.
46. Pinero, J., et al., *The DisGeNET knowledge platform for disease genomics: 2019 update*. *Nucleic Acids Res*, 2020. **48**(D1): p. D845-D855.
47. Tang, Z., et al., *GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses*. *Nucleic Acids Res*, 2017. **45**(W1): p. W98-W102.
48. Gorgoulis, V., et al., *Cellular Senescence: Defining a Path Forward*. *Cell*, 2019. **179**(4): p. 813-827.
49. Pennisi, R., P. Ascenzi, and A. di Masi, *Hsp90: A New Player in DNA Repair?* *Biomolecules*, 2015. **5**(4): p. 2589-618.
50. Pawlowska, E., J. Szczepanska, and J. Blasiak, *DNA2-An Important Player in DNA Damage Response or Just Another DNA Maintenance Protein?* *Int J Mol Sci*, 2017. **18**(7).
51. Liu, Y., et al., *Identification of Hub Genes and Key Pathways Associated With Bipolar Disorder Based on Weighted Gene Co-expression Network Analysis*. *Front Physiol*, 2019. **10**: p. 1081.
52. Nikitaki, Z., *Επιδιόρθωση οξειδωτικών ομαδοποιημένων βλαβών DNA ύστερα από έκθεση καρκινικών κυττάρων σε ιοντίζουσες ακτινοβολίες*. 2017, National Technical University of Athens.
53. McHugh, D. and J. Gil, *Senescence and aging: Causes, consequences, and therapeutic avenues*. *J Cell Biol*, 2018. **217**(1): p. 65-77.

