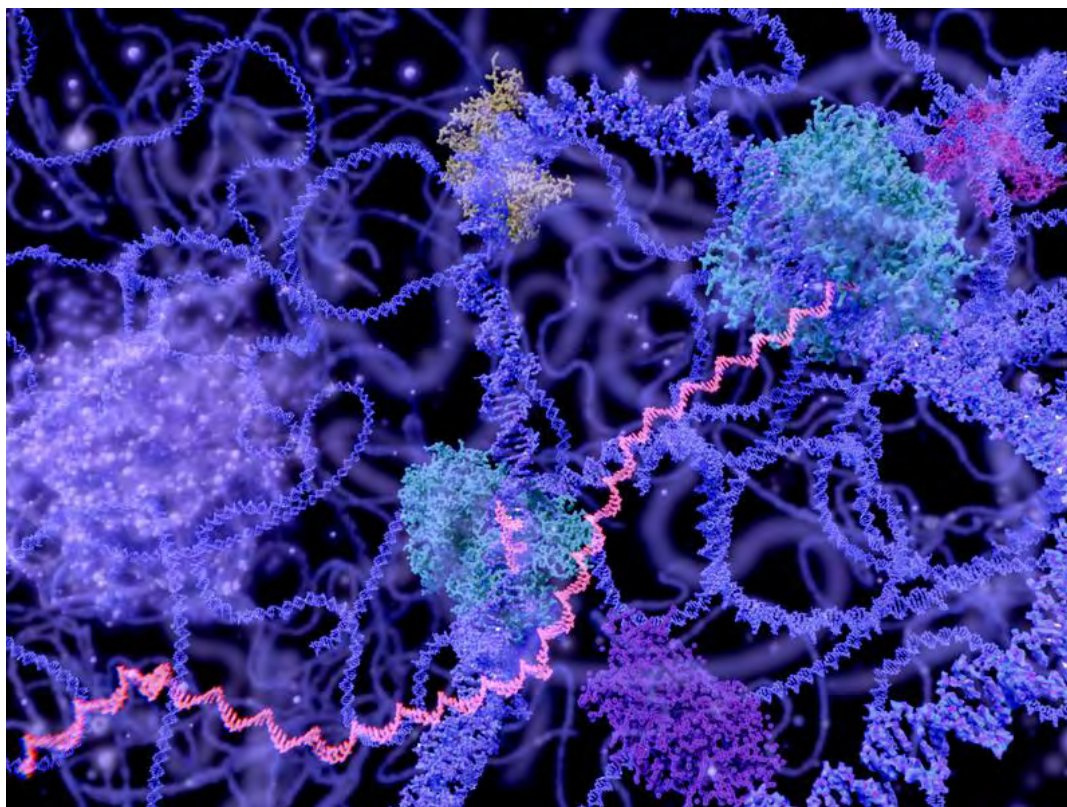


**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ  
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ  
ΕΡΓΑΣΤΗΡΙΟ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ ΚΑΙ ΓΟΝΙΔΙΩΜΑΤΙΚΗΣ**

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**

**ΧΑΡΑΚΤΗΡΙΣΜΟΣ ΤΩΝ lncRNAs CENTRIC1 ΚΑΙ CENTRIC2 ΣΤΟΝ  
ΚΑΡΚΙΝΟ ΤΟΥ ΣΤΟΜΑΧΟΥ**



**Τσιτσιανόπουλος Μάρκος**

**Επιβλέπων καθηγητής:  
Γιακουντής Αντώνιος**

**Λάρισα 2020**

## **Τριμελής Επιτροπή:**

**Γιακουντής Αντώνης:** Επίκουρος Καθηγητής Μοριακής Βιολογίας-Γονιδιωματικής, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας.

**Μαθιόπουλος Κωνσταντίνος:** Καθηγητής Μοριακής Βιολογίας, Τμήμα Βιοχημείας και Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας.

**Λιάκος Παναγιώτης:** Επίκουρος Καθηγητής Ιατρικής Βιοχημείας, Τμήμα Ιατρικής, Πανεπιστήμιο Θεσσαλίας.

## Περιεχόμενα

Περίληψη: .....	5
1. ΕΙΣΑΓΩΓΗ .....	7
1.1. Ο καρκίνος του στομάχου .....	8
1.1.1. Αντιμετώπιση, διάγνωση και θεραπεία .....	8
1.1.2. Η αιτιολογική βάση του καρκίνου του στομάχου .....	9
1.1.3. Κατηγοριοποίηση του καρκίνου του στομάχου .....	10
1.2. Παραλλαγές αριθμού αντιγράφων (Copy Number Variations - CNVs) .....	18
1.2.1. Χαρακτηριστικά των CNVs .....	19
1.2.2. Επίδραση των CNVs στο φαινότυπο .....	21
1.2.3. Ασθένειες συσχετιζόμενες με CNVs .....	22
1.2.4. Ο ρόλος των CNVs στον καρκίνο .....	23
1.2.5. Ο ρόλος των CNVs στον καρκίνο του στομάχου .....	25
1.3. Τα μακρά μη κωδικοποιητικά μόρια RNA (long-non coding RNAs) .....	28
1.3.1. Χαρακτηριστικά και κατηγοριοποίηση των lncRNAs .....	28
2. ΕΡΕΥΝΗΤΙΚΗ ΥΠΟΘΕΣΗ ΚΑΙ ΣΚΟΠΟΣ .....	36
3. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ .....	38
3.1. Υλικά .....	39
3.1.1. Αντιδραστήρια .....	39
3.1.2. Πλασμιδιακοί φορείς .....	39
3.1.3. Κυτταρικές σειρές .....	40
3.2. Μέθοδοι .....	40
3.2.1. Βιοπληροφορική ανάλυση .....	40
3.2.2. Μετασηματισμός βακτηριακών στελεχών και απομόνωση πλασμιδιακού DNA .....	40
3.2.3. Κλωνοποίηση ενθέματος στον pSiCORmod-PGK-puro .....	41
3.2.4. Διαγνωστικές πέψεις πλασμιδιακών φορέων (Diagnostic plasmid digestions) .....	43
3.2.5. Αλληλούχηση .....	43
3.2.6. Υποκυτταρική Κλασμάτωση .....	43
3.2.7. Απομόνωση RNA (RNA Extraction) .....	44
3.2.7. Κατεργασία με DNάση (DNase treatment) .....	44
3.2.8. Σύνθεση cDNA (cDNA synthesis - Reverse Transcription) .....	45
3.2.9. Ποσοτική PCR πραγματικού χρόνου (Real time quantitative PCR - qPCR) .....	46
3.2.10. Πείραμα ελέγχου μεταγραφικής σταθερότητας (Transcriptional stability assay) .....	46
3.2.11. Δοκιμασία ελέγχου σχηματισμού αποικιών (Colony Formation Assay - CFA) .....	47
4. ΑΠΟΤΕΛΕΣΜΑΤΑ .....	48

4.1 Βιοπληροφορική Ανάλυση .....	49
4.1.1 Εντοπισμός γονιδίων με παραλλαγές αριθμού αντιγράφων .....	49
4.1.1 Επαλήθευση κωδικοποιητικού δυναμικού .....	50
4.1.2 Ανάλυση του προφίλ έκφρασης του lncRNA CENTRIC2 .....	52
4.2 Πειραματικά αποτελέσματα .....	57
4.2.1. 1 <sup>ο</sup> Πειραματικό μέρος .....	57
4.2.2. 2 <sup>ο</sup> Πειραματικό μέρος .....	61
5. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ .....	65
6. ΠΑΡΑΡΤΗΜΑ.....	70
6.1 Μηχανισμοί σχηματισμού των CNVs .....	71
Βιβλιογραφία .....	81

## Περίληψη:

Ο καρκίνος του στομάχου αποτελεί μία από τις πιο θανατηφόρες μορφές καρκίνου παγκοσμίως. Συγκαταλέγεται τέταρτος σε συχνότητα εμφάνισης και τρίτος όσον αφορά τη θνησιμότητα στον πληθυσμό, με υψηλά ποσοστά να σημειώνονται στις ανεπτυγμένες χώρες της Ασίας και της Νότιας Αμερικής. Η κύρια αιτία της κακοήθους φύσης του έγκειται στην διάγνωση της νόσου σε προχωρημένα στάδια, καθώς και στη δυσμενή της πρόγνωση, καθιστώντας απαραίτητη την ανάπτυξη ευαίσθητων και εξειδικευμένων μοριακών βιοδεικτών, καθώς και αποτελεσματικών θεραπευτικών προσεγγίσεων. Τα τελευταία χρόνια, τα lncRNAs βρίσκονται στο προσκήνιο της έρευνας στο πεδίο της γονιδιακής έκφρασης ως ρυθμιστικά μόρια, με συμμετοχή σε ένα ευρύ φάσμα βιολογικών διαδικασιών. Μεταλλάξεις σε μόρια αυτού του είδους μεταβάλλουν το προφίλ έκφρασης και τις φυσιολογικές τους λειτουργίες, οδηγώντας στην ανάπτυξη ασθενειών, όπως και στην περίπτωση του γαστρικού καρκίνου. Το ιστοειδικό πρότυπο της έκφρασής τους και η ικανότητα ανίχνευσής τους με μη επεμβατικές μεθόδους τα καθιστά κατάλληλους βιοδείκτες και πολλά υποσχόμενους θεραπευτικούς στόχους. Βασικό χαρακτηριστικό του γαστρικού καρκίνου αποτελεί το φαινόμενο της γονιδιωματικής αστάθειας. Εντοπίζεται σε ένα ιδιαίτερα μεγάλο ποσοστό και συνδέεται στενά με τον εντερικό κατά Lauren ιστολογικό υπότυπο. Χαρακτηρίζεται από υψηλή συχνότητα μεταλλάξεων στο γονιδίωμα ή ακόμα και στο πλαίσιο μιας κυτταρικής γενεαλογίας. Οι μεταλλάξεις αυτές μπορεί να σχετίζονται με μεταβολές στην αλληλουχία των νουκλεϊκών οξέων, με επικρατέστερες τις παραλλαγές στον αριθμό των γονιδιακών αντιγράφων, γνωστές και ως Copy Number Variations. Τα Copy Number Variations αποτελούν παράγοντα γονιδιωματικής αστάθειας, με κύριο χαρακτηριστικό τη μεταβολή της δομής των χρωμοσωμάτων και κατ' επέκταση του αριθμού των γονιδιακών αντιγράφων. Περιλαμβάνουν ελλείψεις, ενθέσεις και διπλασιασμούς αλληλουχιών, οδηγώντας σε μεταβολή της γονιδιακής δόσης σε λειτουργικά, δομικά, καθώς και σε ρυθμιστικά στοιχεία του γονιδιώματος. Αύξηση ή μείωση της γονιδιακής δόσης σε σχέση με την φυσιολογική διπλοειδή μορφή, ενδέχεται να μεταβάλλει την λειτουργία των γονιδίων στόχων και να οδηγήσει σε καρκινογένεση, μέσω της ενεργοποίησης ογκογόνων μονοπατιών σηματοδότησης. Στην παρούσα διπλωματική εργασία, γίνεται απόπειρα ταυτοποίησης και χαρακτηρισμού νέων μορίων lncRNA, τα οποία εμφανίζουν ενίσχυση των γονιδιακών τους αντιγράφων στον καρκίνο του στομάχου. Πραγματοποιούνται ερευνητικές προσεγγίσεις με τη βοήθεια βιοπληροφορικών αναλύσεων και πειραματικών μεθόδων για την μελέτη τους. Παράλληλα, κατασκευάζονται τα απαραίτητα εργαλεία καταστολής της έκφρασής τους με τη βοήθεια μεθόδων γενετικής μηχανικής, με σκοπό την αποσαφήνιση του ρόλου και της συμβολής τους στη γαστρική καρκινογένεση.

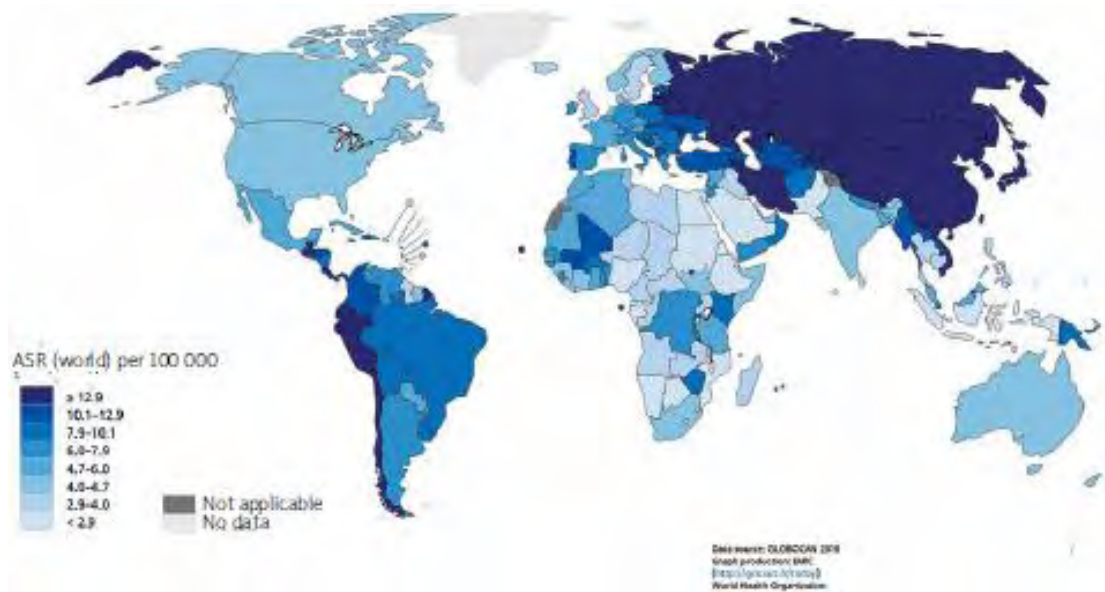
## **Abstract:**

Gastric cancer is considered as one of the deadliest forms of cancer worldwide. Currently ranked fourth most common cancer type and third leading cause of cancer related mortality, the disease notes high rates of incidence occurring in developed countries in Asia and North America. The major cause of its malignant nature is due to its late diagnosis in advanced stages and its poor prognosis, therefore making the development of sensitive and highly specific molecular biomarkers, as well as efficient therapeutic approaches, of great importance. Over the last years, lncRNAs are in the spotlight of gene regulation research field, as regulatory molecules contributing in a broad spectrum of biological processes. Mutations in such molecules alter their expression profile and normal functions, leading to the development of diseases such as gastric cancer. The tissue-specific pattern of their expression and the ability of detecting them by non-invasive methods make them suitable biomarkers and promising therapeutic targets. A key feature of gastric cancer is the phenomenon of genomic instability. It is located in a very high percentage and is closely related to the intestinal Lauren histological subtype. It is characterized by a high frequency of mutations in the genome or even within a cell lineage. Such mutations can be accounted to alterations in the DNA sequence, predominantly in the form of variants in the number of gene copies, also known as Copy Number Variations. Copy Number Variations are a major factor of genomic instability, with the main feature being the change in the structure of chromosomes and consequently the number of gene copies. They involve deletions, insertions and duplications, leading to changes in gene dosage of functional, structural, and regulatory elements of the genome. Increasing or decreasing the gene dosage relatively to the normal diploid form, may alter the function of the target genes and lead to carcinogenesis by activating oncogenic signaling pathways. In the present dissertation, an attempt is made to identify and characterize new lncRNA molecules, whose gene copies are genetically amplified in gastric cancer. Research approaches are carried out with the help of bioinformatic analysis tools and experimental methods for their study. Furthermore, knockdown tools are developed with the help of genetic engineering protocols, in order to clarify their role and contribution to gastric carcinogenesis.

# 1.ΕΙΣΑΓΩΓΗ

### 1.1.Ο καρκίνος του στομάχου

Ο καρκίνος του στομάχου, ή γαστρικός καρκίνος, συγκαταλέγεται ως η τέταρτη κατά σειρά σε συχνότητα, και η τρίτη σε θνησιμότητα μορφή καρκίνου παγκοσμίως, με κυριότερες εστίες εμφάνισης την Ανατολική Ασία και Ευρώπη, καθώς και την Νότια Αμερική, με ποσοστό θνησιμότητας που αγγίζει το 78%<sup>1</sup> (**Εικόνα 1**). Το σύνολο των γαστρικών καρκίνων περιλαμβάνει μια ιδιαίτερα ετερογενή ομάδα κακοθών επιθηλιακών αλλοιώσεων, που επάγονται από ποικιλία προδιαθέσεων και αιτιολογικών παραγόντων, οι οποίοι συμβάλλουν στην επιθετικότητα τους<sup>2,3</sup>. Η πλειοψηφία των περιπτώσεων γαστρικού καρκίνου αναφέρεται σε αδενοκαρκινώματα (95%), με τα πλακώδη (squamous), αδενοπλακώδη (adenosquamous), αδιαφοροποίητα (undifferentiated) και μυελώδη (medullary) καρκινώματα να συμπληρώνουν τη σχετική λίστα<sup>4</sup>. Η πλέον τυπική ηλικία για διάγνωση ανέρχεται προσεγγιστικά στα 60 χρόνια, με τη συχνότητα των περιστατικών να εμφανίζεται διπλάσια στους άνδρες απ' ότι στις γυναίκες<sup>5,6</sup>. Τα συμπτώματα και το προσδόκιμο ζωής διαφέρουν ανάλογα με την εθνικότητα και το υπόβαθρο των ασθενών<sup>7</sup>. Αξιοσημείωτη πληροφορία και ιδιαίτερο χαρακτηριστικό της νόσου αποτελεί η διάγνωση της σε αρκετά προχωρημένο στάδιο, εξαιτίας της έλλειψης ικανών διαγνωστικών βιοδεικτών και άλλων μοριακών εργαλείων<sup>8</sup>. Ως εκ τούτου η πρόγνωση για τον ασθενή είναι εξαιρετικά δυσμενής ενώ το προσδόκιμο επιβίωσης έως 5 έτη από την διάγνωση να υπολογίζεται σε μόλις 32%, καθιστώντας αναγκαία την εύρεση και ανάπτυξη νέων εξειδικευμένων μοριακών εργαλείων, με στόχο την έγκαιρη διάγνωση και την αποτελεσματική θεραπεία της νόσου<sup>9</sup>.



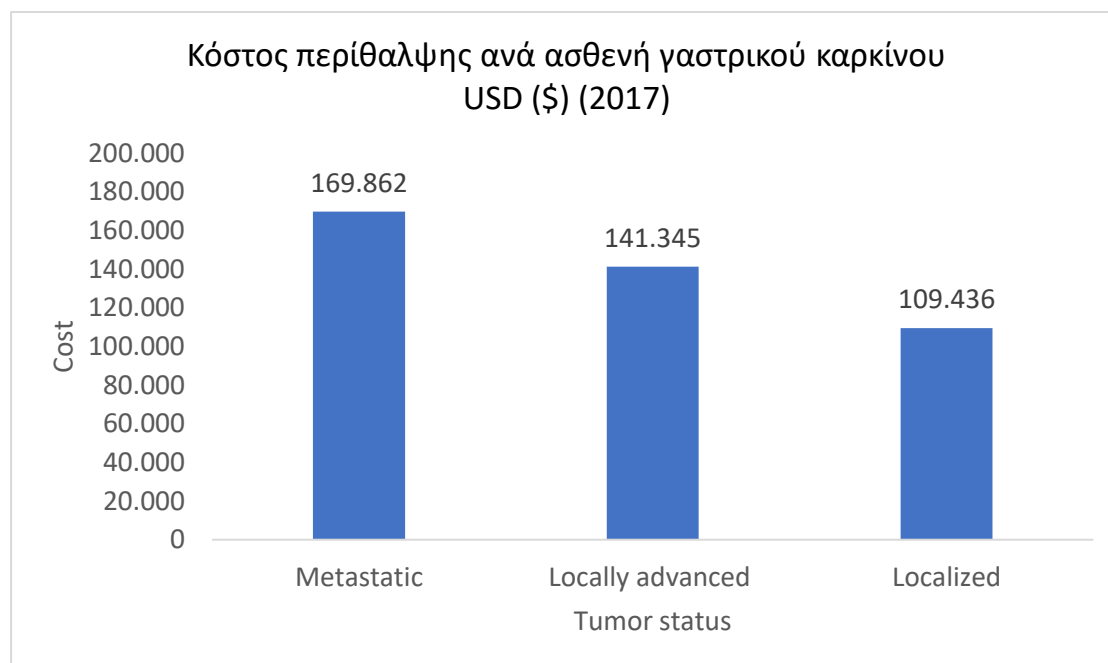
**Εικόνα 1:** Ο χάρτης απεικονίζει τα εκτιμώμενα ποσοστά συχνότητας εμφάνισης για τον καρκίνο του στομάχου παγκοσμίως το 2018 και για τα δύο φύλα, καθώς και για όλες τις ηλικιακές ομάδες<sup>10,11</sup>.

#### 1.1.1. Αντιμετώπιση, διάγνωση και θεραπεία

Η διάγνωση του γαστρικού καρκίνου πραγματοποιείται σε προχωρημένα στάδια, μέσω ενδοσκόπησης ή με τη λήψη βιοψίας από τον ιστό του ασθενή, μεθοδολογία που δεν μπορεί να εφαρμοστεί ως διαγνωστική μέθοδος ρουτίνας για τον γενικό πληθυσμό. Η αντιμετώπιση και η θεραπεία της νόσου προϋποθέτει τη συμμετοχή πολλών επιστημονικών τομέων και εξαρτάται από το στάδιο του όγκου, την ένταση της θεραπευτικής αγωγής, την απόκριση του ασθενούς και τον διαθέσιμο εξοπλισμό. Η αποτελεσματικότερη μέχρι στιγμής



θεραπευτική προσέγγιση περιλαμβάνει χειρουργική αφαίρεσή του όγκου, σε συνδυασμό με χημειοθεραπεία και χρήση ακτινοβολίας για την ύφεση του<sup>1</sup>. Εξαιτίας της καθυστερημένης διάγνωσης του, το κόστος για την περίθαλψη των ασθενών είναι ιδιαίτερα υψηλό, αγγίζοντας έως και τα 160.000 \$ σε περιπτώσεις μεταστατικού καρκίνου<sup>12</sup>. Γι' αυτό το λόγο, κρίνεται αναγκαία η εύρεση και η άμεση εφαρμογή νέων μοριακών εργαλείων με σκοπό την έγκαιρη διάγνωση και την οικονομικά βιώσιμη και αποτελεσματικότερη θεραπεία. (Εικόνα 2).



**Εικόνα 2:** Δεδομένα του μέσου κόστους περίθαλψης του γαστρικού καρκίνου ανάλογα με το στάδιο του όγκου σε ασθενείς άνω των 65 ετών<sup>12</sup>.

### 1.1.2. Η αιτιολογική βάση του καρκίνου του στομάχου

Ο καρκίνος του στομάχου αποτελεί πολυπαραγοντική νόσο, η οποία περιλαμβάνει τόσο γενετικούς, όσο και περιβαλλοντικούς παράγοντες. Πολλοί από αυτούς διαδραματίζουν καθοριστικό ρόλο στην ανάπτυξη του, μεταξύ αυτών οι κληρονομικοί γενετικοί παράγοντες, η λοίμωξη από το ελικοβακτηρίδιο (*Helicobacter pylori*), οι διατροφικοί παράγοντες και ο τρόπος ζωής του εκάστοτε ασθενούς<sup>13-15</sup>.

#### 1.1.2.1. Κληρονομικοί και γενετικοί παράγοντες

Παρόλο που ο κληρονομικός γαστρικός καρκίνος χαρακτηρίζεται από μικρή συχνότητα εμφάνισης, γενετικές μεταβολές έχουν συχνά καταγραφεί σε περιπτώσεις σποραδικών μορφών καρκίνου (στους οποίους γονίδια αλληλεπιδρούν με περιβαλλοντικούς παράγοντες)<sup>16</sup>. Έχει διατυπωθεί η άποψη ότι οι γενετικοί παράγοντες παίζουν καθοριστικό ρόλο στη γαστρική καρκινογένεση, πιθανώς μέσω της επίδρασής τους σε φλεγμονώδεις και ανοσολογικές αποκρίσεις (κυρίως στις λοιμώξεις από το *H.pylori*), αυξάνοντας την ευαισθησία έναντι της νόσου<sup>17</sup>. Ωστόσο, ο αριθμός γονιδίων που εμφανίζουν υψηλή διεισδυτικότητα (δηλαδή υψηλή πιθανότητα έκφρασης καρκινικού φαινοτύπου) για ανάπτυξη καρκίνου του στομάχου είναι μικρός<sup>18</sup>. Ιδιαίτερα αντιπροσωπευτικό παράδειγμα αποτελεί το γονίδιο *CDH1*, το οποίο κωδικοποιεί την E-καδερίνη, στην περίπτωση του

διάχυτου τύπου γαστρικού καρκίνου (diffuse gastric cancer subtype)<sup>19</sup>, ο οποίος θα αναλυθεί περαιτέρω σε επόμενη ενότητα. Ένα ακόμη γονίδιο, αυτό της ιντερλευκίνης 1β (*IL-1β*) έχει αναγνωριστεί ως καίριο, εξαιτίας τη συμβολής του στην έναρξη και ενίσχυση της φλεγμονώδους απόκρισης. Πολυμορφισμοί στο *IL-1β* καθώς και στον ανταγωνιστή του υποδοχέα της ιντερλευκίνης 1 έχουν συσχετιστεί με κίνδυνο εμφάνισης καρκίνου του στομάχου<sup>20</sup>. Επιπλέον, στη λίστα των πολυμορφισμών που σχετίζονται με μια ποικιλία υποτύπων γαστρικού καρκίνου, ανήκουν μονονουκλεοτιδικοί πολυμορφισμοί στα γονίδια *MUC1*, *PSCA* και *PLCE1*<sup>21</sup>.

### 1.1.3. Κατηγοριοποίηση του καρκίνου του στομάχου

Ο καρκίνος του στομάχου διακρίνεται σε δύο κύριες κατηγορίες: στη μορφή πρώιμων (**early**) και προχωρημένων (**advanced**) σταδίων<sup>4</sup>. Το υψηλότερο ποσοστό πρώιμης ανίχνευσης καρκίνου του στομάχου έχει αναφερθεί σε χώρες όπως η Ιαπωνία (50%) και η Κορέα (46-67%). Στην Ευρώπη το αντίστοιχο ποσοστό είναι σημαντικά χαμηλότερο (15%)<sup>22</sup>. Η πρόγνωση στηρίζεται σε μεγάλο βαθμό στο στάδιο ανάπτυξης του όγκου. Το ποσοστό πενταετούς επιβίωσης στην περίπτωση του πρώιμου γαστρικού καρκίνου ανέρχεται σε 85 – 100%, έναντι του προχωρημένου, με ποσοστό 5 – 20%, ενώ ρόλο στην κατηγοριοποίηση παίζει και ο βαθμός μετάστασης σε άλλους ιστούς<sup>23</sup>. Η πρώιμη μορφή ορίζεται ως ένα σύνολο αλλοιώσεων, περιορισμένων στην περιοχή του βλεννογόνου και του υποβλεννογόνου, ανεξαρτήτως διήθησης στους λεμφαδένες. Η προχωρημένη μορφή ορίζεται στα πλαίσια της μεταστατικής ικανότητας και της ενδοσκοπικής παρατήρησης (π.χ. εμφάνιση πολυποειδών αλλοιώσεων, συνοδευόμενων από καλά ή όχι καθορισμένα έλκη, ή διηθητική διάχυτη μορφή χωρίς ενδείξεις όγκων ή έλκους)<sup>24</sup>.

#### 1.1.3.1. Ανατομική κατηγοριοποίηση του καρκίνου του στομάχου

Η θέση του όγκου υπαγορεύει την ανατομική ταξινόμηση του γαστρικού καρκίνου σε εγγύς (**cardia**) ή απομακρυσμένους (**non-cardia**) από την καρδιά του στομάχου<sup>24</sup>. Τα εγγύς και τα απομακρυσμένα αδενοκαρκινώματα εμφανίζονται με ποικίλα κλινικά και επιδημιολογικά χαρακτηριστικά, με τους απομακρυσμένους να παρουσιάζουν μειωμένα ποσοστά τις τελευταίες δεκαετίες, σε αντίθεση με τους εγγύς<sup>25</sup>. Τα εγγύς αδενοκαρκινώματα επηρεάζουν κυρίως τους καυκάσιους πληθυσμούς, ενώ συνοδεύονται συνήθως από γαστροοισοφαγική παλινδρόμηση. Επιπλέον, χαρακτηρίζονται ως επιθετικά, διαπερνώντας τα τοιχώματα του στομάχου και του οισοφάγου, με τελικό προορισμό τους λεμφαδένες<sup>26</sup>. Τα απομακρυσμένα αδενοκαρκινώματα συσχετίζονται σε μεγαλύτερο ποσοστό με το *H.pylori* και αντιπροσωπεύουν περισσότερο από το 60% των περιπτώσεων παγκοσμίως<sup>27</sup>.

#### 1.1.3.2. Ιστολογική κατηγοριοποίηση του καρκίνου του στομάχου

Όπως έχει ειπωθεί σε προηγούμενη ενότητα, η πλειοψηφία των καρκίνων του στομάχου αφορά αδενοκαρκινώματα, τα οποία ωστόσο εμφανίζουν μεγάλη ετερογένεια όσον αφορά την ανάπτυξη, την αρχιτεκτονική, την κυτταρική διαφοροποίηση και το μοριακό τους υπόβαθρο<sup>4,24</sup>. Η ποικιλία αυτή εξηγεί και την ποικιλία ως προς την ιστοπαθολογική κατηγοριοποίησή τους. Οι πλέον επικρατέστερες ιστολογικές κατηγοριοποιήσεις είναι η κατηγοριοποίηση κατά Lauren και η κατηγοριοποίηση σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (WHO).

##### 1.1.3.2.1. Κατηγοριοποίηση κατά Lauren<sup>28</sup>

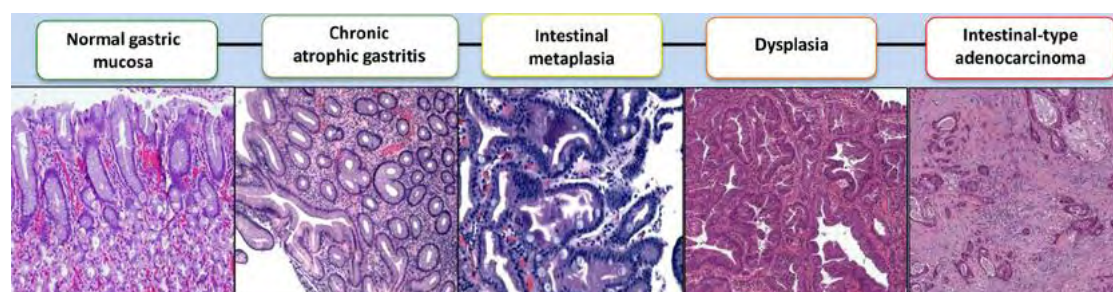
Σύμφωνα με τα κριτήρια του Lauren, ο καρκίνος του στομάχου διακρίνεται σε δύο κύριους τύπους: τον εντερικό (**Intestinal**) και τον διάχυτο (**Diffuse**) τύπο, οι οποίοι

παρουσιάζουν πολυάριθμες διαφορές σε επίπεδο παθολογίας, επιδημιολογίας και αιτιολογίας.

### Εντερικός υπότυπος (Intestinal subtype)

Ο εντερικός υπότυπος αποτελεί τον συνηθέστερο υπότυπο, καθώς αντιστοιχεί στο 54% των περιστατικών στα πλαίσια της διάγνωσης σε πληθυσμούς με υψηλό κίνδυνο εκδήλωσης της νόσου<sup>29</sup>. Εμφανίζεται σε ασθενείς χωρίς οικογενειακό ιστορικό ή μεταβιβάσιμες μεταλλάξεις<sup>30,31</sup>. Διαγιγνώσκεται σε άτομα μεγάλης ηλικίας, η πλειονότητα των οποίων αφορά άντρες, ενώ οι όγκοι διαθέτουν την μορφή εξελκωμένων μαζών<sup>32,33</sup>. Η ογκογένεση σχετίζεται σε μεγάλο βαθμό με λοίμωξη από το *H.pylori*.

Η χρόνια λοίμωξη από το *H.pylori* οδηγεί σε διαδοχικές ιστολογικές αλλοιώσεις, οι οποίες έχουν ως τελικό αποτέλεσμα τον κακοήγη μετασχηματισμό<sup>34,35</sup>. Σύμφωνα με το συγκεκριμένο μοντέλο, η καρκινική προδιάθεση του ιστού ξεκινά ως αποτέλεσμα παρατεταμένης μόλυνσης από το βακτήριο, η οποία ακολουθείται από μη атροφική γαστρίτιδα. Η атροφική γαστρίτιδα μετατρέπεται σε πολυεστιακή атροφική γαστρίτιδα απουσία μεταπλασίας, στη συνέχεια σε εντερική μεταπλασία και τελικά στην εμφάνιση δυσπλασίας<sup>36-39</sup>. Η παρατεταμένη φλεγμονή του γαστρικού ιστού από χρόνια λοίμωξη από το *H.pylori*, είναι πιθανό να προκαλέσει βλάβη στον επιθηλιακό ιστό, η οποία οδηγεί σε атροφία των αδένων, καθώς και σε γαστρική атροφία μέσω απώλειας των θεμέλιων και των τοιχωματικών κυττάρων<sup>40,41</sup>. Το γαστρικό επιθήλιο αντικαθίσταται από εντερική μεταπλασία σε όλη την έκταση του στομάχου, αλλά ιδιαίτερα στην περιοχή του ελάσσονος τόξου. Οι εστίες ανάπτυξης μεταπλασίας θα αποτελέσουν στη συνέχεια τις εστίες της δυσπλασίας που θα ακολουθήσει. Με το πέρασμα του χρόνου, οι εστίες της δυσπλασίας αποκτούν την ικανότητα εισβολής στον χαλαρό συνδετικό ιστό, φαινόμενο που ορίζει το καρκίνωμα ως επιθετικό<sup>42,43</sup> (Εικόνα 3).



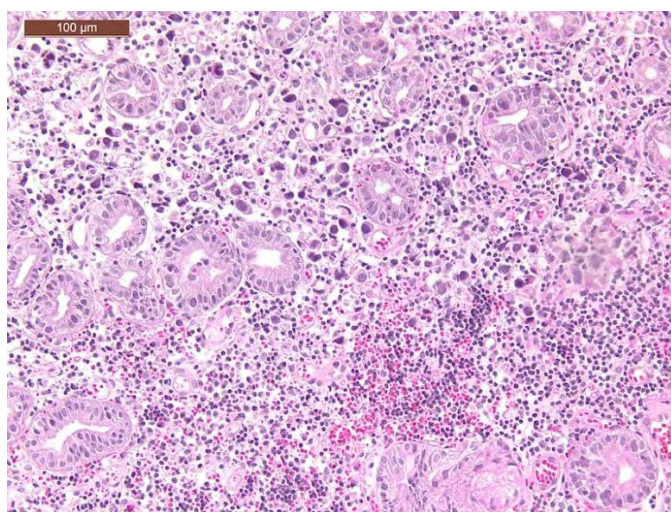
Εικόνα 3: Στάδια ανάπτυξης εντερικού υπότυπου γαστρικού αδενοκαρκινώματος<sup>44</sup> Πηγή: Correa's cascade<sup>36,45</sup>

Αρκετά ογκογονίδια παρουσιάζονται να είναι υπερεκφρασμένα στους όγκους εντερικού υπότυπου. Παρόλα αυτά, καθένα από αυτά τα γονίδια φαίνεται να είναι παρόν σε ξεχωριστά στάδια της νόσου<sup>46</sup>. Για παράδειγμα, το ογκογονίδιο *c-met* εμπλέκεται στην ανάπτυξη περίπου του 20% των καρκίνων γαστρικού εντερικού υπότυπου και η έκφραση του σχετίζεται με προχωρημένα στάδια γαστρικού καρκίνου<sup>47</sup>. Για την ακρίβεια, ορισμένα στελέχη *H.pylori* παράγουν μια πρωτεΐνη τελεστή, η οποία εμφανίζεται να ρυθμίζει μονοπάτια μεταγωγής σήματος που έχουν ως αφετηρία τον υποδοχέα του *c-met* και ταυτόχρονα να είναι υποψήφια για τη συμμετοχή τους στα στάδια της έναρξης και της εξέλιξης της καρκινογένεσης<sup>48,49</sup>. Μεταλλάξεις στο γονίδιο *KRAS* έχουν επίσης εντοπιστεί στα στάδια της εντερικής μεταπλασίας, της δυσπλασίας και των επιθετικών καρκινωμάτων<sup>50,51</sup>.

Μεταλλάξεις σε ογκοκατασταλτικά γονίδια εντοπίζονται σε πολλές περιπτώσεις εντερικού υποτύπου γαστρικού καρκίνου<sup>52</sup>. Όπως αναφέρθηκε σε προηγούμενη ενότητα, το γονίδιο *TP53* που κωδικοποιεί την πρωτεΐνη p53 αποτελεί έναν σημαντικό ρυθμιστή του κυτταρικού κύκλου. Η απώλεια της ετεροζυγωτίας ή η απενεργοποίηση του γονιδίου μέσω μετάλλαξης λαμβάνει χώρα σε ποσοστό μεγαλύτερο του 60% του συνόλου των περιστατικών επιθετικής μορφής γαστρικού καρκίνου<sup>53</sup>. Μη φυσιολογικά επίπεδα έκφρασης του *TP53* έχουν επίσης παρατηρηθεί σε περιπτώσεις χρόνιας γαστρίτιδας, μεταπλασίας και δυσπλασίας. Όσον αφορά το στάδιο της χρόνιας γαστρίτιδας, δεν έχει αποσαφηνιστεί οποιαδήποτε σχέση μεταξύ του *TP53* και του *H.pylori*<sup>54-56</sup>. Επιπλέον παράδειγμα μετάλλαξης σε ογκοκατασταλτικό γονίδιο αποτελεί και η περίπτωση του γονιδίου *APC* τόσο στον εντερικό, όσο και στον διάχυτο υπότυπο<sup>57,58</sup>.

#### Διάχυτος υπότυπος (Diffuse subtype)

Ο διάχυτος υπότυπος καρκίνου εμφανίζεται να είναι πιο επιθετικός από τον εντερικό, διαγιγνώσκεται σε νεότερους σε ηλικία ασθενείς, με ίση συχνότητα σε άνδρες και γυναίκες<sup>33</sup>. Οι όγκοι που συγκαταλέγονται στον συγκεκριμένο υπότυπο, έχουν την τάση να εισβάλλουν στα τοιχώματα του στομάχου και σε γειτονικές δομές, συμπεριλαμβανομένων του οισοφάγου και του δωδεκαδακτύλου, στους οποίους δεν παρατηρείται σχηματισμός αδένων. Ορισμένα παραδείγματα διάχυτης διήθησης του τοιχώματος του στομάχου έχουν παρουσιαστεί να οδηγούν σε νόσο του Brinton, ενώ ο ίδιος υπότυπος συνοδεύεται συνήθως από δομές σφραγιστήρος δακτυλίου (signet-ring)<sup>59</sup> (**Εικόνα 4**).



**Εικόνα 4:** Παράδειγμα απεικόνισης διάχυτου υποτύπου γαστρικού καρκίνου στο μικροσκόπιο. Πηγή: Weisenberg E. Diffuse (poorly cohesive) type adenocarcinoma. PathologyOutlines.com website. <http://www.pathologyoutlines.com/topic/stomachdiffuse.html>.

Οι όγκοι του διάχυτου υποτύπου γαστρικού καρκίνου διαθέτουν έναν σαφή μηχανισμό καρκινογένεσης μέσω ελαττωμάτων στην διαδικασία της διακυτταρικής προσκόλλησης. Η πλειοψηφία των περιπτώσεων προκαλείται από απώλεια έκφρασης της E-καδερίνης, μιας πρωτεΐνης, η οποία διαδραματίζει καθοριστικό ρόλο στην ανώτερη διαδικασία<sup>60,61</sup>. Το γονίδιο *CDH1* κωδικοποιεί μια ομοδιμερή διαμεμβρανική πρωτεΐνη η οποία συντελεί στην συναρμολόγηση του κυττάρου, συμμετέχοντας στο σύμπλοκο κυτταρικής προσκόλλησης (**Cell Adhesion Complex**)<sup>62,63</sup>. Το γονίδιο φαίνεται να έχει ογκοκατασταλτική δράση, καθώς η απενεργοποίηση του προϋποθέτει ένα μοντέλο «δύο

χτυπημάτων»<sup>61</sup>. Η κληρονομική μορφή του διάχυτου γαστρικού καρκίνου (**Hereditary Diffuse Gastric Cancer - HDGC**) φέρει αυτοσωμικό επικρατή τύπο κληρονόμησης. Η νόσος εκδηλώνεται νωρίς κατά τη διάρκεια της ζωής του ασθενούς και έχει πιθανότητα ανάπτυξης γαστρικού καρκίνου μεγαλύτερη από το 80% μέχρι την ηλικία των 80 ετών<sup>64</sup>. Ασθενείς με HDGC συχνά παρουσιάζουν πολυεστιακό (**multifocal**) πρότυπο ογκογένεσης σε μια άθικτη περιοχή του βλεννογόνου, καθιστώντας δυσμενή την διάγνωση της ασθένειας. Τα σποραδικά καρκινώματα διάχυτου υπότυπου έχουν επίσης συσχετιστεί με μεταλλάξεις, απώλεια ετεροζυγωτίας, καθώς και υπερμεθυλίωση του υποκινητή του γονιδίου *CDH1*<sup>65-67</sup>. Ασθενείς που φέρουν μεταλλάξεις στα κύτταρα της γαμετικής σειράς και έχουν οικογενειακό ιστορικό γαστρικού καρκίνου παραπέμπονται για προληπτική γαστρεκτομή<sup>65</sup>.

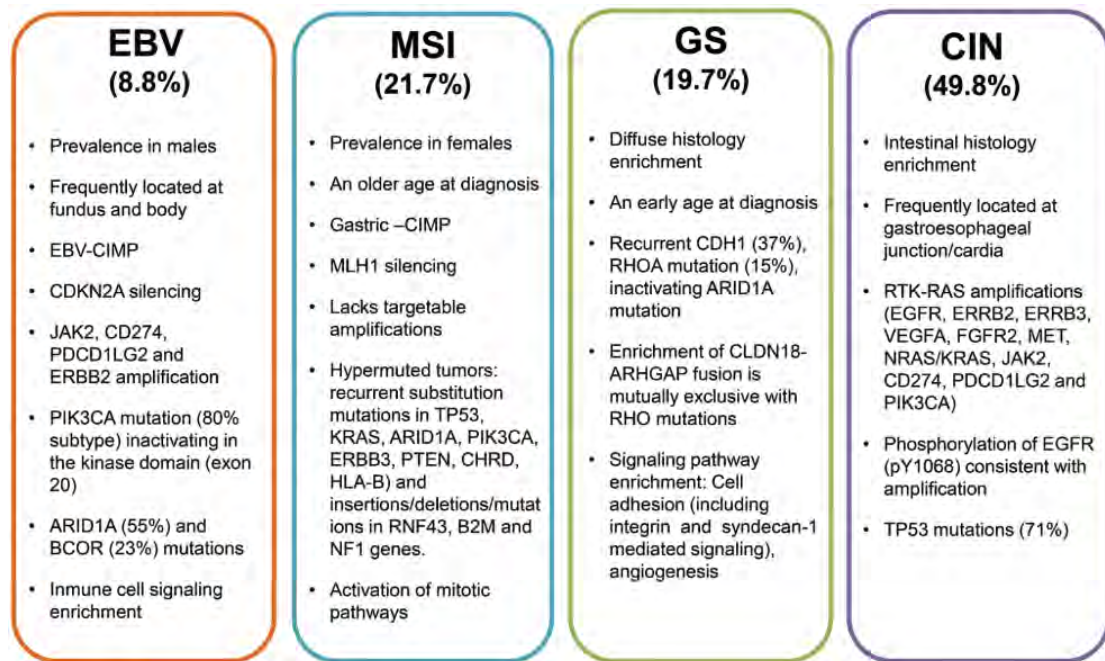
#### 1.1.3.3. Μοριακή κατηγοριοποίηση του καρκίνου του στομάχου

Στο παρελθόν, το σύνολο των γαστρικών όγκων συγκροτούσε ένα ομοιογενές φάσμα ασθενειών, ωστόσο στη σύγχρονη εποχή έχει επικρατήσει ο διαχωρισμός του σε υποκατηγορίες βασισμένες στο φάσμα των γενετικών μεταλλαγών που έχουν ταυτοποιηθεί από την μελέτη του καρκινικού γονιδιώματος. Οι συγκεκριμένες μεταβολές έχουν συνδεθεί με σχετικά κλινικά χαρακτηριστικά, όπως η αιτιολογία το φύλο, η ηλικία της διάγνωσης και ο ανατομικός εντοπισμός, η θεραπευτική αγωγή καθώς και το στάδιο και η επιθετικότητα του όγκου. Σύμφωνα με έρευνα της ερευνητικής κοινοπραξίας **The Cancer Genome Atlas (TCGA)** και από αποτελέσματα που προέκυψαν από 295 δείγματα από γαστρικούς όγκους και καλά χαρακτηρισμένους υπότυπους, έχουν προκύψει τέσσερις βασικοί μοριακοί υπότυποι του καρκίνου του στομάχου<sup>68,69</sup> (**Εικόνα 5,6**):

**EBV<sup>+</sup> υπότυπος (υπότυπος θετικός στον ιό Epstein-Barr)**: Περίπου το 9-10% των γαστρικών αδενοκαρκινωμάτων χαρακτηρίζονται θετικά στον ιό Epstein-Barr<sup>69</sup>. Συνοδεύεται κατά κανόνα από ιδιαίτερες χρωμοσωμικές ανωμαλίες, ένα μοναδικό μοτίβο μεταγραφικής ενεργότητας, καθώς και από τροποποιήσεις σε περιοχές υποκινητών ογκοκατασταλτικών γονιδίων. Η πλειοψηφία των περιπτώσεων περιλαμβάνει άρρενες και οι όγκοι εντοπίζονται στην περιοχή του σώματος ή του θόλου του στομάχου<sup>70</sup>. Ο συγκεκριμένος τύπος γαστρικού καρκίνου φαίνεται να σχετίζεται με ασυνήθιστη τάση για μετάσταση στους λεμφαδένες καθώς και με ελαττωμένα ποσοστά θνησιμότητας σε διαφορετικούς πληθυσμούς<sup>71</sup>. Προσεγγιστικά, το 80% των όγκων που χαρακτηρίζονται θετικοί στον EBV φέρουν μεταλλάξεις στο γονίδιο *PIK3CA*<sup>69</sup> το οποίο κωδικοποιεί την πρωτεΐνη P13Kα<sup>72</sup>. Το συγκεκριμένο γονίδιο συγκαταλέγεται δεύτερο σε συχνότητα εμφάνισης μεταλλάξεων σε ένα ευρύ φάσμα τύπων καρκίνου<sup>73</sup>. Επιπλέον, έχει εντοπιστεί σε ποσοστό μεγαλύτερο από 10% όσον αφορά τις περιπτώσεις που έχουν εξεταστεί από τον TCGA. Μεταλλάξεις στο γονίδιο *PIK3CA* έχουν εντοπιστεί στο 3-42% των περιπτώσεων με EBV<sup>+</sup> και των υπόλοιπων μοριακών υπότυπων γαστρικού καρκίνου<sup>44</sup>. Εξαιρετικό ενδιαφέρον έχει το ερώτημα εάν οι αναστολείς του μονοπατιού σηματοδότησης P13K μπορούν να αποτελέσουν εργαλείο διάκρισης μεταξύ των EBV<sup>+</sup> και των EBV<sup>-</sup> καρκίνων του στομάχου. Μέχρι σήμερα, το σύνολο των EBV<sup>+</sup> περιστατικών έχει παρουσιάσει το μεγαλύτερο ποσοστό υπερμεθυλίωσης σε περιοχές του DNA (**DNA Hypermethylation**) στον κατάλογο των καρκίνων που μελετήθηκαν από το TCGA. Επιπλέον, υπερμεθυλίωση του υποκινητή του γονιδίου *CDKN2A*, που χαρακτηρίζεται ως ένα υψίστης σημασίας ογκοκατασταλτικό γονίδιο στον καρκίνο του στομάχου, έχει ανιχνευθεί σε όλα τα θετικά στον EBV δείγματα<sup>74</sup>. Γονιδιωματικές ενισχύσεις (**amplifications**) στα γονίδια *JAK2*, *CD274*, *PDCD1LG2* και *ERBB2* αλλά και απαισιόφες (**deletions**) στα γονίδια *PTEN*, *SMAD4*, *CDKN2A* και *ARID1A* έχουν επίσης εντοπιστεί στον εν λόγω μοριακό υπότυπο<sup>75</sup>. Η σηματοδότηση JAK/STAT διαδραματίζει καθοριστικό ρόλο στον κυτταρικό πολλαπλασιασμό, στην κυτταρική διαφοροποίηση και στην απόπτωση και έχει

αποτελέσει αντικείμενο εντατικής έρευνας στον καρκίνο του στομάχου<sup>76</sup>. Αυξημένη έκφραση του γονιδίου PD-L1, έχει συσχετιστεί με δυσμενή πρόγνωση στις περιπτώσεις γαστρικού καρκίνου, συμβάλλοντας πιθανώς σε καταστολή και αποφυγή του ανοσοποιητικού<sup>77,78</sup>. Αντίστοιχα, αναστολείς των πρωτεϊνών PI3K και του JAK2, και ανταγωνιστές της PD-L1/2 ίσως έχουν θεραπευτική αξία για τις περιπτώσεις EBV<sup>+</sup> γαστρικού καρκίνου.

**Αστάθειας μικροδορυφόρων (Microsatellite instability - MSI):** Ο συγκεκριμένος μοριακός υπότυπος αποτελεί έναν από τους κυριότερους μοριακούς τύπους γονιδιωματικής αστάθειας (δηλαδή υψηλής συχνότητας μεταλλάξεων στο γονιδίωμα μιας κυτταρικής γενεαλογίας) που έχουν συσχετιστεί με θηλυκούς ασθενείς σχετικά μεγάλης ηλικίας. Αναφέρεται σε γενετικές μεταλλαγές που προκαλούνται εξαιτίας της απενεργοποίησης επιδιορθωτικών ενζύμων. Σχετίζεται με τον εντερικό υπότυπο γαστρικού αδενοκαρκινώματος κατά Lauren με χαμηλά επίπεδα μετάστασης στους λεμφαδένες και καλύτερα ποσοστά πρόγνωσης<sup>79-81</sup>. Οι περιπτώσεις αστάθειας μικροδορυφόρων χαρακτηρίζονται κατά κανόνα από συσσώρευση μεταλλάξεων στα γονίδια *PIK3CA*, *ERBB3*, *ERBB2* και *EGFR*, σε συνδυασμό με μεταλλάξεις που παρουσιάζονται σε θέσεις οι οποίες έχουν ταυτοποιηθεί και σε άλλα είδη καρκίνων<sup>69,80</sup>. Προηγούμενες μελέτες υποστηρίζουν ότι η κύρια αιτία πρόκλησης αστάθειας μικροδορυφόρων στον καρκίνο του στομάχου περιλαμβάνει υπερμεθυλίωση του υποκινητή του γονιδίου *MLH1*, το οποίο εμπλέκεται στο σύστημα επιδιόρθωσης αταίριαστων ζευγών βάσεων στο DNA (**DNA mismatch repair - MMR**)<sup>82,83</sup>. Ανάλογα με το επίπεδο της αστάθειας, τα περιστατικά μπορούν να διακριθούν ως υψηλής (**MSI-H**) ή χαμηλής (**MSI-L**) συχνότητας<sup>82</sup>. Αντίθετα, όγκοι που δεν εμφανίζουν αστάθεια σε κανένα γενετικό τόπο μικροδορυφόρων κατηγοριοποιούνται ως μικροδορυφορικά σταθεροί (**microsatellite stable tumors - MSS**)<sup>82</sup>. Ο φαινότυπος υψηλής συχνότητας έχει συσχετιστεί σε μεγάλο βαθμό με τα αδενοκαρκινώματα εντερικού υποτύπου και δύναται καλύτερης πρόγνωσης από τους μικροδορυφορικά σταθερούς όγκους και τους φαινότυπους χαμηλής συχνότητας<sup>69,80,82,84</sup>. Έχει επίσης αναπτυχθεί η υπόθεση ότι η μικροδορυφορική αστάθεια παίζει καθοριστικό ρόλο στην μεταλλαγή γονιδίων σχετικών με τον έλεγχο του κυτταρικού κύκλου (π.χ. *TGFBR2*, *EPHB2*, *E2F4*), την επιδιόρθωση βλαβών του DNA (*MRE1*, *ATR*) και την αποπτωτική σηματοδότηση (π.χ. *BAX*)<sup>82,85,86</sup>.



Εικόνα 5: Τα κύρια χαρακτηριστικά και οι σημαντικότερες γενετικές αλλοιώσεις σε σχέση με τον εκάστοτε μοριακό υπότυπο, σύμφωνα με το TCGA.<sup>44</sup>

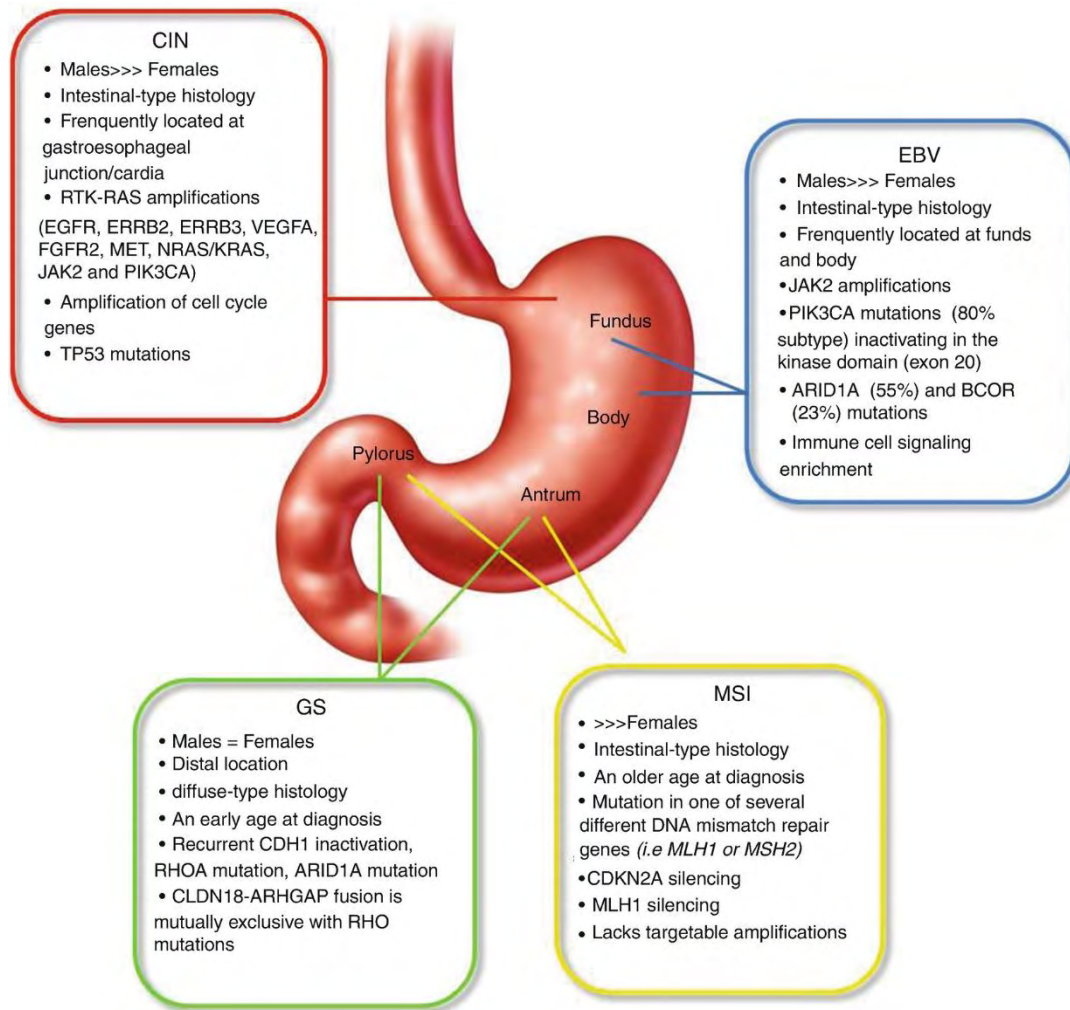
**Γονιδιωματικής σταθερότητας (Genomically stable - GS):** Ο συγκεκριμένος υπότυπος αντιπροσωπεύει το 20% των περιστατικών καρκίνου του στομάχου που μελετήθηκαν από τον TCGA και σχετίζεται στενά με τον διάχυτο υπότυπο γαστρικού αδενοκαρκινώματος<sup>69</sup>. Διαγιγνώσκεται κυρίως σε ασθενείς νεότερης ηλικίας, χαρακτηρίζεται από μικρή συχνότητα μεταλλάξεων στο γονίδιο *TP53*, καθώς και από χαμηλό βαθμό ανευπλοειδίας. Μεταλλάξεις στα γονίδια *CDH1* και *RHOA* έχουν εντοπιστεί σε ποσοστό 13-37% και 14-25% αντίστοιχα, σε περιπτώσεις διάχυτου υπότυπου καρκίνου του στομάχου, επισημαίνοντας τις επιπτώσεις της αλλαγής στους μηχανισμούς κυτταρικής προσκόλλησης για τον συγκεκριμένο ιστολογικό υπότυπο<sup>69,87,88</sup>. Όπως έχει προαναφερθεί, μεταλλάξεις του γονιδίου *CDH1* στα κύτταρα της γαμετικής σειράς σχετίζονται με HDGC, καθώς και με δυσμενή φαινότυπο και κλινικά αποτελέσματα. Αντίθετα, μεταλλάξεις του γονιδίου *RHOA* οδηγούν σε αλληλεπιδράσεις της κωδικοποιούσας πρωτεΐνης με άλλες πρωτεΐνες στα κύτταρα, με αποτέλεσμα την επαγωγή μορφολογικών μεταβολών και μηχανισμών κυτταρικής μετανάστευσης, λειτουργίες καίριες για την ανάπτυξη και την αύξηση του μεγέθους του όγκου<sup>89-91</sup>. Μια επιπλέον γενετική μεταβολή που αφορά τη λειτουργία του *RHOA* περιλαμβάνει το γονίδιο *ARHGAP26*, το οποίο κωδικοποιεί για μια G πρωτεΐνη, με ρόλο θετικού ρυθμιστή για το προϊόν του γονιδίου *RHOA*. Έχουν καταγραφεί περιπτώσεις χρωμοσωμικής μετάθεσης μεταξύ των γονιδίων *CLDN18* και *ARHGAP26*, με το χιμαιρικό προϊόν να επηρεάζει τη ρύθμιση της πρωτεΐνης *RHOA*. Η συγκεκριμένη περίπτωση γονιδιακής σύντηξης δύναται να επηρεάσει και το αλληλόμορφο αγρίου τύπου του *CLDN18*, το οποίο κωδικοποιεί μια πρωτεΐνη με ρόλο στενοσυνδέσμου (**tight junction protein**), που εκφράζεται σε υψηλά επίπεδα στο επιθήλιο του στομάχου και επικουρεί την διαδικασία της διακυτταρικής προσκόλλησης<sup>92</sup>. Αξίζει να αναφερθεί ότι η σύντηξη μεταξύ *CLDN18* και *ARHGAP26*, είναι αμοιβαία αποκλειστική με μεταλλάξεις στο γονίδιο *RHOA*. Επομένως, μεταλλαγές στα μονοπάτια που αφορούν κυτταρική προσκόλληση στον συγκεκριμένο μοριακό υπότυπο είναι πιθανό να καταστήσουν δυνατή την ταυτοποίηση νέων θεραπευτικών στόχων και συγκεκριμένα των συμμετεχόντων στο μονοπάτι δράσης της πρωτεΐνης *RHOA*<sup>44</sup>.

**Χρωμοσωμικά ασταθείς (Chromosomally unstable - CIN):** Το 50% των περιστατικών καρκίνου του στομάχου χαρακτηρίζονται ως χρωμοσωμικά ασταθείς, με τον συγκεκριμένο φαινότυπο να αφορά κατά κανόνα την περιοχή της καρδίας του στομάχου. Το 71% των γαστρικών όγκων αυτού του υποτύπου φέρουν μεταλλάξεις στο γονίδιο *TP53*, καθώς και στα γονίδια *ARID1A*, *KRAS*, *PIK3CA*, *RNF43*, *ERBB2* και *APC*<sup>44</sup>. Ο CIN γαστρικός καρκίνος, φαίνεται να σχετίζεται με τον εντερικό ιστολογικό υπότυπο κατά Lauren, παρουσιάζοντας παράλληλα αυξημένα ποσά έκφρασης της p53 πρωτεΐνης, η απώλεια της οποίας συνεπάγεται και τα υψηλά ποσοστά ανευπλοειδίας (δηλαδή του μη φυσιολογικού αριθμού χρωμοσωμάτων, ή τμημάτων αυτών, στο κύτταρο)<sup>93</sup>. Οι παραπάνω πληροφορίες συμβαδίζουν επίσης με τις παρατηρήσεις που αφορούν γονίδια με τη μέγιστη συχνότητα απώλειας ετεροζυγωτίας, όπως οι γενετικοί τόποι των *APC* (36%) και *TP53* (33%). Μεταλλαγές στο γονίδιο *TP53*, έχουν συσχετιστεί με προκαρκινικές γαστρικές αλλοιώσεις, υποδεικνύοντας ότι η απώλεια της λειτουργίας της p53 προδιαθέτει για γαστρική καρκινογένεση<sup>94</sup>. Επιπλέον, αυξημένα επίπεδα στους χρωμοσωμικά ασταθείς γαστρικούς καρκίνους παρουσιάζει και η συχνότητα φωσφορυλίωσης του υποδοχέα EGFR, η οποία είναι ανάλογη της αυξημένης έκφρασής του.

Ωστόσο, το πιο σημαντικό χαρακτηριστικό του χρωμοσωμικά ασταθούς γαστρικού καρκίνου αφορά την υψηλή συχνότητα μεταβολής των αντιγράφων σε διάσπαρτους γενετικούς τόπους του γονιδιώματος (**Copy Number Variations - CNVs**)<sup>95</sup>. Χαρακτηριστικό παράδειγμα αποτελούν γονίδια που κωδικοποιούν υποδοχείς με δράση κινάσης τυροσίνης (*RTKs*) και ορισμένων προσδετών (πχ *VEGFA*), όπου αύξηση των αντιγράφων τους οδηγεί σε μη φυσιολογική κυτταρική ανάπτυξη<sup>96</sup>. Ενισχύσεις γενετικών τόπων στον χρωμοσωμικά ασταθή υπότυπο, αφορούν και γονίδια που συμμετέχουν στον κυτταρικό κύκλο (π.χ. *CCNE1*, *CCND1*, and *CDK6*)<sup>97</sup>. Το φαινόμενο της αύξησης της έκφρασης εξαιτίας του αυξημένου αριθμού αντιγράφων αφορά και μεταγραφικούς παράγοντες, όπως οι *KLF5*, *GATA4* και *GATA6* οι οποίοι ενισχύονται σε αντίγραφα στον καρκίνο του στομάχου<sup>98</sup>. Έχει αποδειχθεί ότι οι τρεις αυτοί μεταγραφικοί παράγοντες συμμετέχουν στο ίδιο σύμπλοκο και λειτουργούν συνεργατικά ως ογκογόνοι ρυθμιστές κυτταρικών γενεαλογιών (**lineage survival oncogenes**)<sup>99</sup>, επάγοντας τον κυτταρικό πολλαπλασιασμό<sup>100</sup>. Επιπλέον, το 30% των ασθενών με γαστρικό αδενοκαρκίνωμα έχει παρουσιάσει ενίσχυση στα γονίδια των *KLF5*, *GATA4* και *GATA6*, καθιστώντας τους τρεις αυτούς μεταγραφικούς παράγοντες πιθανούς ερευνητικούς στόχους για την περαιτέρω μελέτη της νόσου<sup>101</sup>.

Ο καρκίνος του στομάχου χαρακτηρίζεται από υψηλό ποσοστό γεγονότων γονιδιωματικής και χρωμοσωμικής αστάθειας. Η συχνή εμφάνιση των CNVs στον γαστρικό καρκίνο και η γενική τους τάση να προάγουν την καρκινογένεση, τα καθιστά κατάλληλα για μελέτη και χαρακτηρισμό. Δεδομένης της πληθώρας πληροφοριών που μπορεί να αντληθεί όσον αφορά την ταυτοποίηση σηματοδοτικών μονοπατιών ή κλινικοπαθολογικών στοιχείων, είναι δυνατή η διαδικασία προσδιορισμού και ανάπτυξης νέων βιοδεικτών της ασθένειας<sup>95</sup>. Ως αποτέλεσμα, μέσω του ακριβούς προσδιορισμού και της ανίχνευσης των CNVs και υπό τις συνθήκες ετερογένειας που χαρακτηρίζουν τη νόσο, μπορεί να γίνει δυνατή η ανάπτυξη διαγνωστικών, προγνωστικών και θεραπευτικών προσεγγίσεων του γαστρικού καρκίνου.





**Εικόνα 6:** Επισκόπηση των τεσσάρων μοριακών υποτύπων γαστρικού καρκίνου, σύμφωνα με τον TCGA<sup>102</sup>.

## 1.2. Παραλλαγές αριθμού αντιγράφων (Copy Number Variations - CNVs)

Η γενετική ποικιλομορφία, δηλαδή οι διαφορές στις κωδικές και μη κωδικές περιοχές του γονιδιώματος, είναι το κύριο στοιχείο που καθιστά τον άνθρωπο και κατ' επέκταση κάθε ζωντανό οργανισμό, μια μοναδική οντότητα<sup>103</sup>. Διακρίνονται ποικίλοι τύποι γενετικών μεταλλαγών που συμβάλλουν σε αυτή την ποικιλομορφία, εκ των οποίων ένα μεγάλο ποσοστό παραμένει ανεξερεύνητο. Ο ανθρώπινος πληθυσμός παρουσιάζει εκτεταμένες παραλλαγές (**variants**) σε ένα σύνολο χρωμοσωμικών τμημάτων και στον αριθμό των στοιχείων που περιέχονται εντός αυτών. Οι παραλλαγές διακρίνονται στις εξής δύο κατηγορίες<sup>104</sup>:

**Μονονουκλεοτιδικές παραλλαγές (Single Nucleotide Variants - SNVs):** Μεταβολές σε μεμονωμένες νουκλεοτιδικές θέσεις σε συγκεκριμένα σημεία του γονιδιώματος. Αντιπροσωπευτικό παράδειγμα αποτελούν οι μονονουκλεοτιδικοί πολυμορφισμοί (**Single Nucleotide Polymorphisms - SNPs**), στους οποίους ανήκουν οι παραλλαγές με συχνότητα μεγαλύτερη του 1% στον πληθυσμό.

**Παραλλαγές αριθμού αντιγράφων (Copy Number Variants - CNVs):** Μεταβολές οι οποίες αφορούν αλλαγές στον αριθμό και το μήκος των χρωμοσωμικών τμημάτων, σε σύγκριση με ένα γονιδίωμα αναφοράς. Το μήκος των τμημάτων, όσον αφορά την τάξη μεγέθους, κυμαίνεται από 1kb έως και μερικά Mb<sup>105</sup>.

Τα CNVs εντάσσονται στην κατηγορία των δομικών παραλλαγών (**Structural Variants - SVs**), δηλαδή γονιδιωματικών ανακατατάξεων που επηρεάζουν την αρχιτεκτονική των χρωμοσωμάτων. Τα SVs μπορεί να χαρακτηρίζονται ως ισορροπημένα (**balanced**), δηλαδή χωρίς συνολική απώλεια ή περίσσεια γενετικού υλικού εντός ή μεταξύ των χρωμοσωμάτων, είτε αντίθετα ως μη ισορροπημένα (**unbalanced**)<sup>106</sup>. Οι χρωμοσωμικές ανευπλοειδίες εκφράζουν μια ακραία μορφή μη ισορροπημένης δομικής παραλλαγής. Οι κυριότεροι τύποι των SVs είναι οι εξής<sup>105</sup> (**Εικόνα 7**):

**Μεταθέσεις (Translocations):** Αλλαγές θέσης χρωμοσωμικών τμημάτων μεταξύ δύο ομόλογων ή μη χρωμοσωμάτων, χωρίς απώλεια ή περίσσεια γενετικού υλικού.

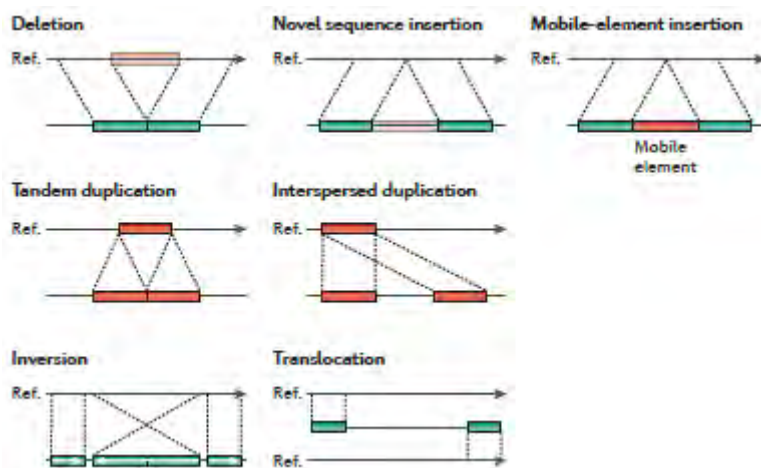
**Αναστροφές (Inversions):** Αλληλουχίες οι οποίες έχουν αντιστραφεί σε σχέση με την αρχική τους κατεύθυνση σε σχέση με το υπόλοιπο χρωμόσωμα. Ορίζονται ως περικεντρικές (**pericentric**), αν εντοπίζονται στην περιοχή του κεντρομερούς ενώ ως παρακεντρικές (**paracentric**) εάν εντοπίζονται σε ποιο απομακρυσμένες χρωμοσωμικές περιοχές.

**Ενθέσεις (Insertions):** Προσθήκες νέων αλληλουχιών στο γονιδίωμα, ως επί το πλείστον μεταθετών στοιχείων (**Mobile Element Insertions – MEIs**).

**Ελλείψεις (Deletions) και διπλασιασμοί (Duplications):** Απαλοιφές ή επαναλήψεις τμημάτων στο DNA (αποκλίνοσες του φυσιολογικού) σε σύγκριση με ένα γονιδίωμα αναφοράς. Οι επαναλήψεις αυτές μπορούν να είναι συνεχείς (**in tandem**) ή να είναι διάσπαρτες στο γονιδίωμα (**interspersed**). Τα CNVs ανήκουν στην πλειοψηφία τους στις δύο αυτές δομικές παραλλαγές.

Περαιτέρω είδη SVs μπορούν να θεωρηθούν τα ισοχρωμοσώματα (**isochromosomes**)<sup>107</sup> όπου τα άκρα δύο μη ομόλογων χρωμοσωμάτων αποτελούν κατοπτρικά είδωλα, οι εύθραυστες χρωμοσωμικές περιοχές (**fragile sites**)<sup>108</sup>, τα δακτυλιοειδή χρωμοσώματα (**ring chromosomes**)<sup>109</sup>, οι τμηματικές μονογονικές δυσωμίες (**segmental uniparental disomies**)<sup>110</sup>, όπου ένα τμήμα χρωμοσώματος έχει κληρονομηθεί σε δύο δόσεις

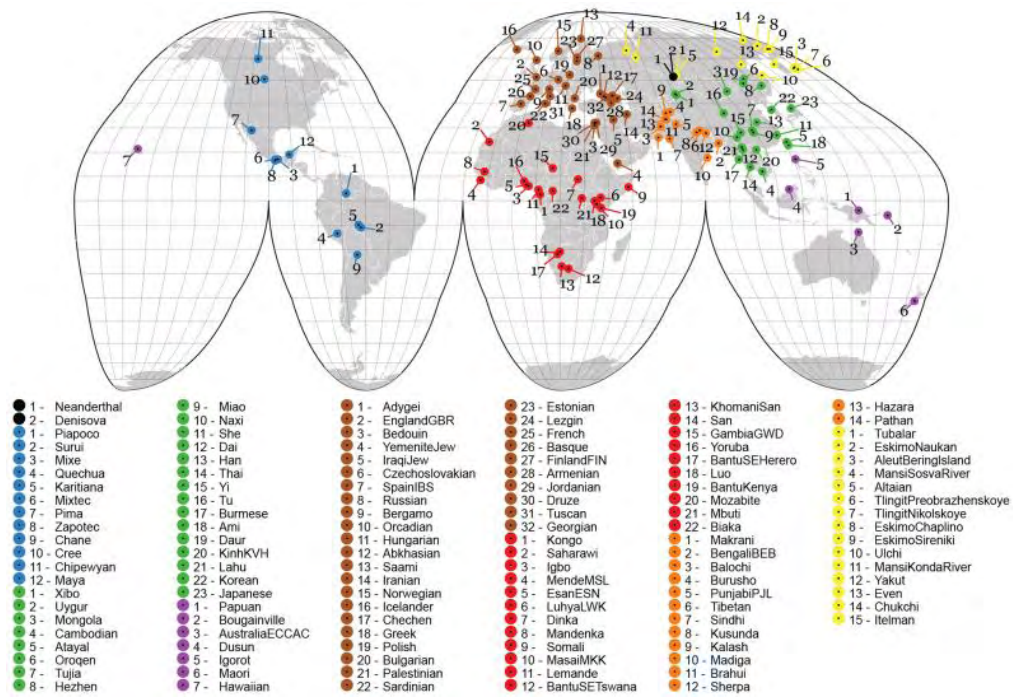
μόνο από ένα γονέα και επιπλέον εξωχρωμοσωμικά τμήματα DNA, γνωστά ως double minutes τα οποία είναι προϊόντα του φαινομένου της χρωμοσωμικής θραύσης (**chromothripsis**)<sup>104</sup>. Κατά τη θραύση λαμβάνει χώρα ένα σύνολο ταυτόχρονων δομικών ανακατατάξεων, που έχουν ως αποτέλεσμα πολλαπλές ελλείψεις, μεταθέσεις και αναστροφές στο γονιδίωμα.



**Εικόνα 7:** Κατηγορίες δομικών παραλλαγών στο γονιδίωμα<sup>105</sup>.

### 1.2.1. Χαρακτηριστικά των CNVs

Στο ανθρώπινο γονιδίωμα υγιών ατόμων, ένα ποσοστό που κυμαίνεται μεταξύ 4,8% με 9,5%, αντιστοιχεί σε CNVs<sup>111</sup>. Ποικιλία στον αριθμό των επαναλαμβανόμενων αλληλουχιών δύναται επίσης να εμφανίζεται σε διαφορετικά όργανα και ιστούς του ίδιου ατόμου<sup>112</sup>. Οι παραλλαγές στον αριθμό των αντιγράφων φαίνεται να είναι τουλάχιστον το ίδιο σημαντικές με τους μονονουκλεοτιδικούς πολυμορφισμούς όσον αφορά την διάκριση μεταξύ των ατόμων στον πληθυσμό<sup>113</sup>. Υπάρχουν παραδείγματα CNVs, ευρέως διαδεδομένων στο γονιδίωμα υγιών ατόμων, που δεν εμπλέκονται στην πρόκληση οποιασδήποτε ασθένειας. Η συστηματική ανάλυση τους μέσω του 1000 Genomes Project και άλλων ερευνών σε πολυπληθείς πληθυσμούς υγιών ατόμων, έχει ρίξει φως στην συχνότητα και τα αποτελέσματα αυτής της ποικιλομορφίας στο ανθρώπινο γονιδίωμα<sup>114</sup> (**Εικόνα 8**). Τυπικό παράδειγμα αποτελεί το γονίδιο της σιαλικής αμυλάσης (*AMY1*), του οποίου ο αριθμός των αντιγράφων είναι ευθέως ανάλογος με την ποσότητα της πρωτεΐνης που ανιχνεύεται στον οργανισμό<sup>115</sup>. Ο μέσος αριθμός των αντιγράφων είναι μεγαλύτερος σε πληθυσμούς με κατανάλωση υψηλών επιπέδων αμύλου σε σχέση με αυτούς που καταναλώνουν λιγότερες ποσότητες<sup>116</sup>. Τα CNVs ευθύνονται για την πλειοψηφία της ποικιλότητας μεταξύ των ατόμων όσον αφορά την έκταση σε ζεύγη βάσεων, καθώς υπερβαίνουν συνδυαστικά τόσο το σύνολο των SNVs όσο και των προσθηκών και ελλείψεων μικρής κλίμακας (**Indels**)<sup>117</sup>.



Εικόνα 8: Ανάλυση της κατανομής των CNVs σε διάφορους πληθυσμούς ανά τον κόσμο<sup>114</sup>.

Έχει επικρατήσει η υπόθεση ότι η κατανομή των CNVs στο γονιδίωμα δεν είναι τυχαία<sup>118</sup>. Έχουν εντοπιστεί αρκετά «θερμά» σημεία (“hot” spots) στα οποία οι παραλλαγές στον αριθμό των αντιγράφων εκδηλώνονται με υψηλότερη συχνότητα<sup>119,120</sup>. Οι περιοχές του γονιδιώματος που φέρουν αυτά τα σημεία είναι γνωστές ως CNVRs (**C**opy **N**umber **V**ariation **R**egions)<sup>121</sup>. Για την ακρίβεια πάνω από 250 περιοχές, επιμέρους μήκους 1Mb, έχουν βρεθεί να φέρουν παραλλαγές σε ποσοστό άνω του 50% της αλληλουχίας τους<sup>122</sup>. Τα CNVs εμφανίζονται ως επί το πλείστον στις περιοχές του κεντρομερούς και των τελομερών, φαινόμενο που εξηγείται εύκολα από τα μεγάλα ποσοστά επαναλήψεων που περιέχουν<sup>123</sup>. Η συσχέτιση μεταξύ των χρωμοσωμάτων και του περιεχομένου τους σε CNVs δεν είναι ξεκάθαρη, καθώς κάποια χρωμοσώματα εμφανίζουν μεγαλύτερη περιεκτικότητα από άλλα. Για παράδειγμα, οι παραλλαγές στον αριθμό των αντιγράφων καταλαμβάνουν περίπου το 19,8% της αλληλουχίας στο χρωμόσωμα 18, ενώ στα χρωμοσώματα 16,17,19,22 και Y, το ποσοστό ξεπερνά το 41%<sup>107,124,125</sup>. Τα παραπάνω δεδομένα αντικατοπτρίζουν τις διαφορές στην δομή του γονιδιώματος μεταξύ των χρωμοσωμάτων, καθιστώντας κάποια από αυτά πιο ευαίσθητα σε μεταλλαγή από άλλα. Επιπλέον, υπάρχουν ενδείξεις συσχέτισης μεταξύ γονιδίων και CNVRs, καθώς περιοχές πλούσιες σε γονίδια είναι ταυτόχρονα πλούσιες σε CNVs και αντίστροφα, μέσω πειραμάτων ανάλυσης του αριθμού των επαναλήψεων, σε συνδυασμό με ανάλυση της πυκνότητας των εξονίων των γονιδίων στόχων<sup>122</sup>.

Εκτός από την μη τυχαία κατανομή των CNVs, φαίνεται να υπάρχουν και ενδείξεις αλληλοεπικάλυψης γονιδιακών αλληλουχιών και επαναλαμβανόμενων αλληλουχιών κατά μήκος του γονιδιώματος<sup>121,126</sup>. Αναλύσεις οντολογίας έχουν υποδείξει έναν μεγάλο αριθμό γονιδίων με παραλλαγές στον αριθμό των αντιγράφων, σχετικών με την ανοσολογική απόκριση και την απόκριση στα εξωτερικά ερεθίσματα<sup>127</sup>. Γονίδια που έχουν ρόλο στην ορθή λειτουργία των αισθήσεων, όπως οσφρητικοί υποδοχείς και φωτοϋποδοχείς, έχουν εντοπιστεί σε CNVRs<sup>121,128,129</sup>. Στην ίδια κατηγορία ανήκουν και γονίδια που συμμετέχουν σε διάφορες νευροφυσιολογικές διεργασίες και στην ανάπτυξη του εγκεφάλου. Πιο συγκεκριμένα, το σύμπλεγμα γονιδίων της πρωτοκαδερίνης στο χρωμόσωμα 5, τα οποία

έχουν ρόλο στην ομαλή εγκεφαλική λειτουργία, εμφανίζει μεγάλη περιεκτικότητα σε CNVs<sup>129</sup>. Τέλος, υπάρχουν και οι περιπτώσεις γονιδίων με πολύ μικρή συχνότητα εμφάνισης παραλλαγών στον αριθμό των αντιγράφων. Τυπικό παράδειγμα αποτελούν γονίδια που κωδικοποιούν πρωτεΐνες με επικράτειες πρόσδεσης στο γονιδίωμα, καθώς και πρωτεΐνες που εμπλέκονται στον μεταβολισμό των νουκλεϊκών οξέων<sup>130</sup>. Ομοίως, γονίδια που συμμετέχουν στην κυτταρική σηματοδότηση και στον πολλαπλασιασμό σημειώνουν μικρά ποσοστά εμφάνισης CNVs<sup>121</sup> στο φυσιολογικό γονιδίωμα. Το γεγονός αυτό αντικατοπτρίζει το μέγεθος της επιλογής που υφίστανται, εξαιτίας του σημαντικού τους ρόλου στην ανάπτυξη και τη ρύθμιση της γονιδιακής έκφρασης<sup>118</sup>.

### 1.2.2. Επίδραση των CNVs στο φαινότυπο

Οι παραλλαγές στον αριθμό των αντιγράφων μπορούν να επηρεάσουν τα γονίδια με πολλούς τρόπους και αυτό αντίστοιχα μπορεί να γίνει αντιληπτό από τις αλλαγές που παρατηρούνται στο φαινότυπο. Η διαφορά στον αριθμό των αντιγράφων ολόκληρων γονιδίων (**gene dosage effect**) δύναται να συνεπάγεται και διαφορά στην γονιδιακή έκφραση, αφού αύξηση ή μείωση του θα προκαλούσε αύξηση ή μείωση των επιπέδων έκφρασης αντίστοιχα<sup>131,132</sup>. Εκτιμάται ότι πάνω από το 17,7% της κληρονομήσιμης μεταβολής στην έκφραση οφείλεται στην ύπαρξη CNVs<sup>133</sup>. Όσον αφορά τα γονίδια που η λειτουργία τους επηρεάζεται από τον αριθμό των αντιγράφων, περιπτώσεις απώλειας (**loss of function**) ή κέρδους λειτουργίας (**gain of function**) μπορούν να οδηγήσουν σε επιβλαβή ή ωφέλιμο φαινότυπο. Η ποικιλομορφία στην έκφραση λόγω των CNVs διαφέρει από γονίδιο σε γονίδιο, όπως φαίνεται στις περιπτώσεις των *GSTT1* και *UGT2B17*, τα οποία συνεισφέρουν στην μεταφορά μορίων στο εσωτερικό του κυττάρου, και του γονιδίου *SNCA*<sup>134,135</sup>. Το συγκεκριμένο γονίδιο κωδικοποιεί την α-συνουκλείνη, η οποία συμβάλλει στην διατήρηση του αριθμού των κυστιδικών μεταφορέων στην συναπτική σχισμή. Το τελευταίο μάλιστα, παρουσιάζει σχεδόν άμεση θετική συσχέτιση μεταξύ των αριθμών των αντιγράφων και των επιπέδων mRNA και παραγόμενης πρωτεΐνης<sup>135</sup>. Γενικά, έχει παρατηρηθεί ότι η πλειοψηφία των συσχετίσεων μεταξύ CNVs και επιπέδων έκφρασης αφορά θετική αναλογία, με την αύξηση του αριθμού των αντιγράφων να συνεπάγεται αύξηση της έκφρασης. Οι περιπτώσεις αρνητικής συσχέτισης ανέρχονται σε ποσοστό 5-15% του συνόλου<sup>133</sup>.

Η αλλαγή στον αριθμό των γονιδιακών αντιγράφων δεν αποτελεί τη μοναδική περίπτωση επίδρασης των CNVs στα επίπεδα της γονιδιακής έκφρασης. Περίπου το 50% των επιδράσεων περιλαμβάνουν διαταραχές κωδικών αλληλουχιών, όπως απαλοιφές εξονίων, ρυθμιστικών στοιχείων και λοιπών λειτουργικών περιοχών. Τα CNVs που εμπλέκουν τμήματα γονιδίων οδηγούν δυνητικά στη δημιουργία παραλλαγμένων πρωτεϊνών, μέσω διαδικασιών ανακατανομής εξονίων (**exon shuffling**)<sup>136</sup>, σχηματισμού προϊόντων εναλλακτικού ματίσματος (**splice variants**)<sup>137</sup> ή ακόμη και χμαιοικικών γονιδίων. Τυπικό παράδειγμα αποτελεί η σύντηξη των κωδικών περιοχών δύο γονιδίων, των *OR51A4* και *OR51A2* που κωδικοποιούν οσφρητικούς υποδοχείς<sup>126</sup>. Ελλείψεις ή διπλασιασμοί εκτός των κωδικών περιοχών των γονιδίων μπορούν επίσης να επηρεάσουν την γονιδιακή έκφραση μέσω μεταβολής της θέσης ή της λειτουργίας καίριων ρυθμιστικών στοιχείων (**position effects**)<sup>138</sup>. Επιπλέον στοιχεία επιβεβαιώνουν την επίδραση των CNVs στην εξ αποστάσεως ρύθμιση, ακόμα και σε αποστάσεις άνω των 2 Mb από γονίδια<sup>133</sup>. Έλλειψη ενός ρυθμιστικού στοιχείου με κατασταλτική λειτουργία μπορεί να προκαλέσει αύξηση της έκφρασης ενός γονιδίου στόχου του. Αντίστοιχα, διπλασιασμοί μεταξύ ενός υποκινητή και του σημείου έναρξης της μεταγραφής (**Transcriptional Start Site - TSS**) μπορεί να οδηγήσουν σε μείωση της έκφρασης του γονιδίου που ρυθμίζει, εξαιτίας της αύξησης της μεταξύ τους απόστασης<sup>139</sup>.

Πιθανό αποτέλεσμα της ύπαρξης παραλλαγών στον αριθμό των αντιγράφων αποτελεί και ο μωσαϊκισμός. Πιο συγκεκριμένα, ένα γονίδιο δύναται να μετατεθεί σε θέση παρακείμενη σε περιοχές ετεροχρωματινής, οι οποίες εμφανίζουν υψηλό αριθμό συμπύκνωσης και άρα μειωμένη έως και μηδενική μεταγραφική ενεργότητα. Η μετάθεση κοντά σε τέτοιες περιοχές μπορεί να προκαλέσει εξάπλωση της ετεροχρωματινικής δομής σε περιοχές ευχρωματινής, οι οποίες παρουσιάζουν χαμηλό αριθμό συμπύκνωσης και υψηλή μεταγραφική ενεργότητα, και άρα σε σίγηση του συγκεκριμένου γονιδίου<sup>140</sup>. Το ποσοστό των ετεροχρωματινικών περιοχών ποικίλει μεταξύ των κυττάρων, με αποτέλεσμα ο μωσαϊκισμός σε επίπεδο γονιδιακής έκφρασης να οφείλεται σε γονίδια που μπορεί να εκφράζονται σε ένα κύτταρο, ενώ σε ένα άλλο όχι. Τα CNVs μπορούν επίσης να οδηγήσουν σε φαινόμενα ιστοειδικότητας όσον αφορά συγκεκριμένες ρυθμιστικές αλληλουχίες ή ισομορφές εξονίων, οδηγώντας σε μεταβολές στην έκφραση των προϊόντων τους, αυξάνοντας ακόμα περισσότερο την πολυπλοκότητα των μηχανισμών που διέπουν την ομαλή λειτουργία των κυττάρων<sup>141</sup>.

Ανάλογα με την κλινική τους σημαντικότητα, τα CNVs μπορούν να διακριθούν σε μία από τις εξής τρεις κατηγορίες<sup>142</sup>:

**Καλοήθειες παραλλαγές (Benign variants):** Σε αυτές ανήκουν τα CNVs που ενισχύονται σε άτομα του πληθυσμού αλλά δεν προδιαθέτουν για παθολογικούς φαινοτύπους. Εμφανίζονται με μεγάλη συχνότητα στον πληθυσμό και είναι υπεύθυνα για την εγκαθίδρυση της ποικιλομορφίας στο εσωτερικό του.

**Παθογόνες παραλλαγές (Pathogenic variants):** Τα CNVs αυτής της κατηγορίας έχουν συσχετιστεί έντονα με ασθένειες και εμφανίζουν υψηλή διεισδυτικότητα.

**Παραλλαγές αμφιβόλου σημαντικότητας (Variants of uncertain significance - VUS):** Σε αυτές εντάσσονται όλες οι παραλλαγές που δεν ορίζονται ως καλοήθειες ή παθογόνες. Η συγκεκριμένη κατηγορία αποτελείται από CNVs τα οποία δεν έχουν ταυτοποιηθεί, ή αν έχουν ταυτοποιηθεί δεν υπάρχουν επαρκή και ξεκάθαρα δεδομένα για την κλινική τους σημαντικότητα. Εντοπίζονται τρεις υποκατηγορίες VUS: Αρχικά, οι πιθανώς παθογόνες (**likely pathogenic**), στις οποίες τα CNVs έχουν καταγραφεί σε μοναδικές περιπτώσεις ασθενών ατόμων και ίσως περιέχουν γονίδια που σχετίζονται με την ασθένεια. Οι πιθανώς καλοήθειες (**likely benign**) όπου τα CNVs δεν περιέχουν κάποιο γονίδιο, ωστόσο το μέγεθός τους ξεπερνά τα όρια του συνηθισμένου. Τέλος, υπάρχουν και οι παραλλαγές αμφιβόλου σημαντικότητας που δεν επιδέχονται υποκατηγοριοποίησης (**VUS with no sub-classification**). Η συγκεκριμένη υποκατηγορία περιλαμβάνει CNVs που περιέχουν γονίδια, για τα οποία ωστόσο δεν υπάρχουν πληροφορίες όσον αφορά τη λειτουργία τους, κι αν αυτή επηρεάζεται από τον αριθμό των αντιγράφων<sup>142</sup>. Κάθε CNV χαρακτηρίζεται αρχικά ως VUS και στη συνέχεια μπορεί να ενταχθεί σε μία από τις άλλες δύο κατηγορίες<sup>143</sup>.

### 1.2.3. Ασθένειες συσχετιζόμενες με CNVs

Παρά την ύπαρξη καλοηθών περιπτώσεων στο ανθρώπινο γονιδίωμα, η επίδραση των CNVs στην γονιδιακή έκφραση, καθώς και οι διαταραχές που μπορούν να προκαλέσουν στην δομή και τη λειτουργία των γονιδίων, τα καθιστούν έναν καίριο αιτιολογικό παράγοντα για έναν μεγάλο αριθμό ασθενειών<sup>118</sup>. Δεδομένου του μεγάλου αριθμού των γονιδίων ο γενετικός τόπος των οποίων υπόκειται σε CNVs, είναι λογικό κάποια από αυτά να συνδέονται με το ευρύ φάσμα παθολογιών που συνδέονται με αλλαγή στον αριθμό αντιγράφων DNA. Περίπου το 50% των γονιδίων με CNVs που έχουν καταχωρηθεί στην βάση δεδομένων OMIM σχετίζονται με μεντελικά νοσήματα<sup>129</sup>. Τη σχετική λίστα συμπληρώνουν γονιδιωματικές

διαταραχές (**Genomic Disorders** - δηλαδή διαταραχές που οφείλονται σε αστάθεια του γονιδιώματος)<sup>104</sup> και λοιπές ασθένειες.

Μεγάλος αριθμός μεταβολών στον αριθμό των αντιγράφων οδηγεί αποδεδειγμένα σε μονογονιδιακές ασθένειες. Κλασικό παράδειγμα αποτελεί η νόσος του Huntington, κατά την οποία αύξηση των τρινουκλεοτιδικών επαναλήψεων CAG στο γονίδιο *HTT* που κωδικοποιεί την χαντικτίνη οδηγεί σε αδυναμία ελέγχου των κινήσεων. Η ασθένεια φέρει αυτοσωμικό επικρατές πρότυπο κληρονόμησης<sup>144</sup>. Όσον αφορά τις ασθένειες με υπολειπόμενο τύπο κληρονόμησης, η δημιουργία ημιζυγωτίας εξαιτίας της έλλειψης ενός γονιδιακού αλληλομόρφου θα μπορούσε να προκαλέσει την εκδήλωση μιας παθολογίας μετάλλαξης στο εναπομείναν αντίγραφο. Αντίστροφα, ο διπλασιασμός ενός φυσιολογικού αλληλομόρφου δύναται να υπερκαλύψει τις επιπτώσεις της ύπαρξης μιας παθολογίας μετάλλαξης στο αντίστοιχο αλληλόμορφο του ομόλογου χρωμοσώματος, σώζοντας με αυτό τον τρόπο τον φαινότυπο. Ο ρόλος των CNVs απαντάται επίσης και στις περιπτώσεις επικρατών νοσημάτων που χαρακτηρίζονται από ποικίλη διεισδυτικότητα<sup>113</sup>. Ενδεικτικό παράδειγμα αποτελεί ο γενετικός τόπος *APP* στο χρωμόσωμα 21, ο οποίος έχει προκύψει από διπλασιασμό του γονιδίου *APP* και σχετίζεται με την νόσο Alzheimer. Η συγκεκριμένη αύξηση στον αριθμό των αντιγράφων ευθύνεται για την αυξημένη αποθήκευση ποσοτήτων αμυλοειδούς στον εγκέφαλο<sup>145</sup>. Ομοίως, τριπλασιασμός του γονιδίου *SNCA*, ο οποίος οδηγεί στον σχηματισμό μεγάλου αριθμού σωματίων Lewy (**Lewy bodies**), έχει συσχετιστεί με νόσο του Parkinson με επικρατή τρόπο κληρονόμησης<sup>146</sup>. Κανένα από τα δύο παραπάνω γονίδια δεν εμφανίζει ανάλογες παραλλαγές σε υγιή άτομα, γεγονός που τα καθιστά πρωταρχικούς αιτιολογικούς παράγοντες για την εκάστοτε ασθένεια.

Η διαθεσιμότητα και η εφαρμογή προηγμένων γονιδιωματικών τεχνολογιών έχει οδηγήσει στην αποτελεσματικότερη ανίχνευση των CNVs και στην ανακάλυψη νέων ασθενειών που προκαλούνται από αυτά. Ως αποτέλεσμα, τα CNVs αναγνωρίζονται όλο και περισσότερο ως πηγή γενετικής ποικιλομορφίας στον παγκόσμιο πληθυσμό, αλλά και ως ένα φορτίο μεταλλάξεων, άρρηκτα συνδεδεμένο με τις ανθρώπινες ασθένειες.

#### 1.2.4. Ο ρόλος των CNVs στον καρκίνο

Οι παραλλαγές στον αριθμό των αντιγράφων τόσο σε επίπεδο χρωμοσωμάτων όσο και σε επίπεδο χρωμοσωμικών τμημάτων παρατηρούνται σε υψηλή συχνότητα στον καρκίνο<sup>147-149</sup>. Ενίσχυση ογκογονιδίων αλλά και ελλείμματα ογκοκατασταλτών έχουν αποδειχθεί καθοριστικά γεγονότα για την πρόοδο της καρκινογένεσης<sup>150</sup>. Παραδείγματα αποτελούν οι ενισχύσεις του γονιδίου *EGFR* στα γλοιώματα<sup>151</sup>, του *MYCN* στο νευροβλάστωμα<sup>152</sup>, του *MYC* στην οξεία μυελώδη λευχαιμία<sup>153</sup> καθώς και του *ERBB2* σε περιπτώσεις καρκίνου του μαστού, των ωθηκών και καρκίνων του πνεύμονα<sup>154</sup>. Ομοίως, ελλείψεις ογκοκατασταλτικών γονιδίων όπως του *PTEN* και του *VHL*, έχουν καταγραφεί σε μια ποικιλία όγκων<sup>155,156</sup>. Πολλές φορές οι όγκοι εξαρτώνται από τέτοια γεγονότα προκειμένου να εξασφαλίσουν τον πολλαπλασιασμό και την επιβίωση, επιλέγοντας έτσι υποπληθυσμούς καρκινικών κυττάρων με τέτοιους γονότυπους που τους δίνουν πλεονέκτημα εξάπλωσης και τελικά κυριαρχίας μέσα στον όγκο. Η στόχευση αυτών των γονιδίων οδηγεί στον περιορισμό ή ακόμα και την εξάλειψή του. Το φαινόμενο αυτό είναι γνωστό ως ογκογονιδιακή εξάρτηση (**oncogene addiction**)<sup>157-160</sup>. Πέρα από την ενίσχυση ογκογονιδίων, έχουν παρατηρηθεί αλλαγές και στον αριθμό των αντιγράφων ποικίλων χρωμοσωμικών περιοχών. Αναλύσεις ολόκληρου του γονιδιώματος έχουν καταγράψει 76.000 αυξήσεις και 55.000 ελλείματα σε ένα σύνολο 3.131 περιστατικών καρκίνου<sup>148</sup>. Μια τυπική περίπτωση καρκίνου χαρακτηρίζεται από 17% ενισχύσεις και 16% ελλείψεις, συγκριτικά με το 0,5% των φαινομένων σε φυσιολογικά δείγματα<sup>148</sup>. Τα παραπάνω δεδομένα

υποδεικνύουν ότι τα CNVs είναι ένα εξαιρετικά συχνό φαινόμενο στα καρκινικά κύτταρα. Επιπλέον, αναλύσεις σε 17 τύπους όγκων αποδεικνύουν ότι το 25% του γονιδιώματος επηρεάζεται από μεταβολές σε επίπεδο ολόκληρων χρωμοσωμάτων ενώ το 10% από μικρότερης κλίμακας χρωμοσωμικές αλλαγές<sup>148</sup>. Οι αλλαγές αυτές περιλαμβάνουν περιοχές μεμονωμένων γονιδίων και χαρακτηρίζονται με τον όρο εστιακά φαινόμενα (**focal events**)<sup>148,161</sup>. Αξιοσημείωτο είναι το γεγονός ότι οι περισσότερες περιοχές που ενισχύονται στον καρκίνο μέσω τέτοιων φαινομένων (focally amplified) περιέχουν ογκογονίδια, ενώ αντίστοιχα οι περιοχές που απαλείφονται (focally deleted) περιέχουν ογκοκατασταλτικά γονίδια<sup>161,162</sup>. Τέτοιου είδους παρατηρήσεις οδηγούν στο συμπέρασμα ότι η επιλογή που ασκείται κατά την ογκογένεση δύναται να προκαλέσει στοχευμένη, και όχι τυχαία, ενίσχυση ή έλλειψη συγκεκριμένων περιοχών στο γονιδίωμα των καρκινικών κυττάρων.

Γονιδιωματικές εστιακές ενισχύσεις (**focal amplifications**) μπορούν επίσης να αφορούν και γονίδια που συμμετέχουν στην απόκριση των φαρμάκων. Περίπου το 10% των περιπτώσεων καρκίνου εμφανίζουν εστιακή ενίσχυση της περιοχής 1q21.2 στο χρωμόσωμα 1, το οποίο περιέχει το αντιαποπτωτικό γονίδιο *MCL1*. Ένα άλλο παράδειγμα αφορά το γονίδιο *BCL2L1* στην περιοχή 20q11.21<sup>163</sup>. Και τα δύο γονίδια είναι απαραίτητα για την επιβίωση του κυττάρου, επομένως η ενίσχυσή τους στο περιβάλλον του όγκου θα μπορούσε να προσδώσει ένα μεγάλο πλεονέκτημα στην επιβίωση των καρκινικών κυττάρων.

Μεταβολές στη δομή των χρωμοσωμάτων σε διάφορες περιοχές μπορούν επίσης να αποτελέσουν αιτία παθογένεσης για ορισμένους τύπους καρκίνου. Στο πολλαπλό μύελωμα (Multiple Myeloma - MM), η πρόοδος της νόσου χαρακτηρίζεται μερικώς από εστιακές ενισχύσεις πλησίον του μικρού βραχίονα του χρωμοσώματος 1 (chr 1q). Μελέτες έχουν ταυτοποιήσει μια περιοχή 10-15 Mb η οποία συμμετέχει στην ενίσχυση της περιοχής 1q12-23 στο MM<sup>164</sup>. Η εν λόγω περιοχή περιέχει έναν μεγάλο αριθμό γονιδίων που έχουν υποστεί ενίσχυση ή απορρύθμιση στην έκφρασή τους. Τυπικά παραδείγματα αποτελούν τα γονίδια *CKS1B*<sup>165,166</sup>, *MUC1*<sup>167</sup> και *MCL1*<sup>168</sup>.

Πολλοί τύποι καρκίνων χαρακτηρίζονται από συσχέτιση μεταξύ γονδιακών ενισχύσεων και αυξημένης ανθεκτικότητας σε αντικαρκινικές θεραπείες<sup>169</sup>. Ασθενείς με καρκίνο των ωθηκών που φέρουν ενίσχυση της περιοχής 1q12-21 εμφανίζουν ανθεκτικότητα σε θεραπεία με την καρκινική ουσία σισπλατίνη<sup>170</sup>. Ενισχύσεις του γονιδίου *CCNE1* είναι παρούσες στο 25% των περιστατικών καρκίνου των ωθηκών και σχετίζονται με μειωμένα ποσοστά επιβίωσης και αντίσταση σε αναστολείς των CDK2<sup>171</sup>. Σε περιπτώσεις καρκίνων του πνεύμονα, παρατηρείται μια αύξηση 11-13 φορές στον αριθμό των αντιγράφων της περιοχής 7q21.12. Η πλειοψηφία των γονιδίων εντός αυτής της περιοχής, συμπεριλαμβανομένων και γονιδίων που κωδικοποιούν μεταφορείς φαρμάκων όπως τα γονίδια *MDR1/ABCB1*, παρουσιάζουν αυξημένη έκφραση<sup>172</sup>. Τα παραπάνω παραδείγματα υποδεικνύουν πως συγκεκριμένες περιοχές στο γονιδίωμα είναι εστιακά ενισχυμένες και συνδέονται με μεταβολή στην έκβαση της ασθένειας και στην απόκριση στη φαρμακευτική αγωγή. Το ενδεχόμενο οι επιλεγμένες δομικές μεταβολές και οι γονδιακές ενισχύσεις να αποτελούν στοχαστικό φαινόμενο ή αποτέλεσμα μιας άρτια δομημένης απόκρισης ενάντια στις θεραπευτικές προσεγγίσεις προϋποθέτει περαιτέρω έρευνα.



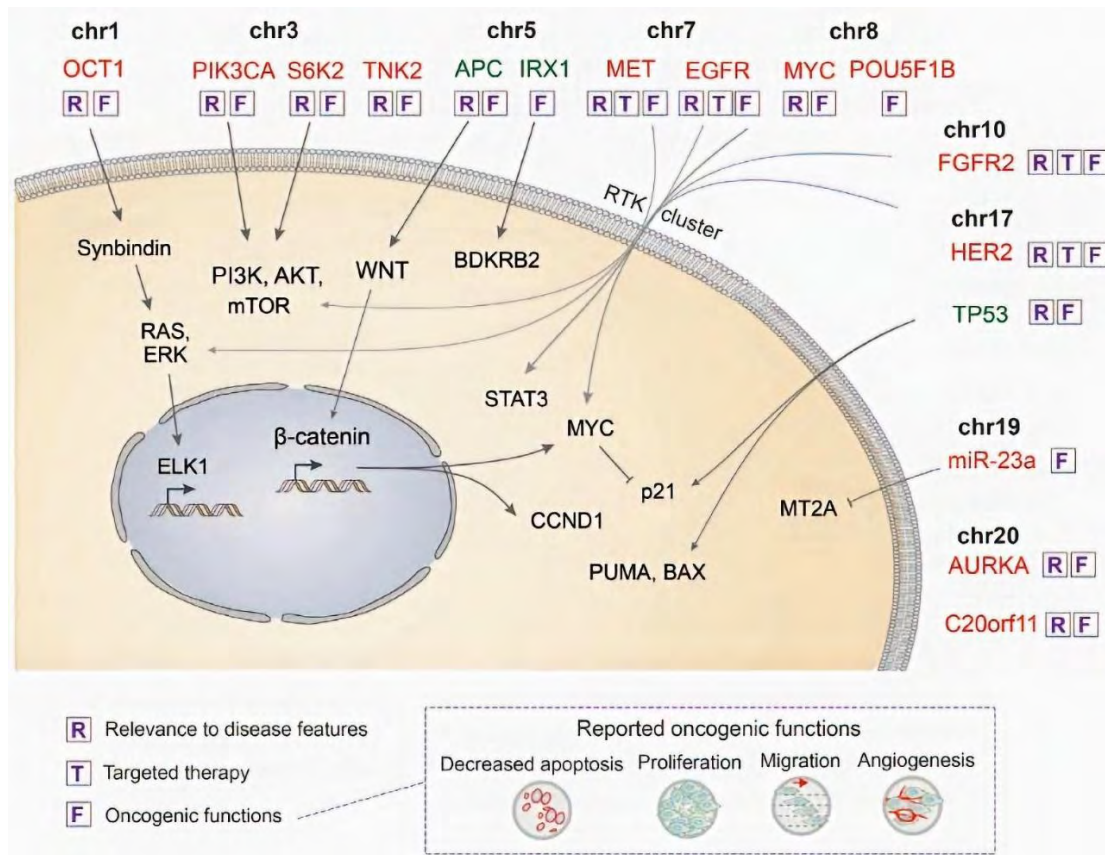
### 1.2.5. Ο ρόλος των CNVs στον καρκίνο του στομάχου

Τα εναρκτήρια στάδια ανάπτυξης και εξέλιξης του γαστρικού καρκίνου χαρακτηρίζονται σε μεγάλο ποσοστό από απορρύθμιση μονοπατιών μεταγωγής σήματος μέσω γενετικών και επιγενετικών αλλαγών<sup>95,173-175</sup>. Αλλαγές στον αριθμό των αντιγράφων δύνανται να επηρεάσουν την έκφραση μιας πληθώρας γονιδίων και να οδηγήσουν στην καρκινογένεση<sup>176</sup>. Τα τελευταία χρόνια, πολλά γονίδια έχουν προταθεί για τον πιθανό ογκοεπαγωγικό ή τον ογκοκατασταλτικό τους ρόλο μέσω της συμμετοχής του σε πολλαπλά στάδια πρωτογενούς και μεταστατικού καρκίνου τα οποία σχετίζονται με ένα ευρύ σύνολο μεταλλάξεων. Έρευνες στο παρελθόν έχουν αναδείξει έναν μεγάλο αριθμό ενθέσεων και ελλείψεων στο DNA των καρκινικών κυττάρων, με τις πρώτες να υπερτερούν έναντι των δεύτερων<sup>176</sup>. Τα CNVs είναι ένας ιδιαίτερα σημαντικός παράγοντας μεταβολής της γονιδιακής έκφρασης και ιδιαίτερα της δράσης διάφορων ογκοεπαγωγικών και ογκοκατασταλτικών γονιδίων, γεγονός που εξηγεί την κλινική τους σημασία στον καρκίνο του στομάχου. Ιδιαίτερα ενδιαφέρουσα είναι η αυξημένη παρουσία ενθέσεων και ελλείψεων και λοιπών CNVs σε ασθενείς με γαστρικό καρκίνο που παρουσιάζει μετάσταση σε λεμφαδένες<sup>95</sup>. Επιπλέον έρευνες έχουν υποδείξει στοχευμένα φαινόμενα δομικών παραλλαγών, με συχνές ενθέσεις στα χρωμοσώματα ή στις χρωμοσωμικές περιοχές 1q, 5p, 7, 8, 13 and 20 και ελλείψεις στα χρωμοσώματα 1p, 3p, 4, 5q, 9p, 17p, 18q, 19p, 21 and 22<sup>87</sup>. Παρόλο που έχει ταυτοποιηθεί ένας μεγάλος αριθμός περιοχών που φέρουν CNVs στον καρκίνο του στομάχου<sup>177</sup>, η μελέτη και ο χαρακτηρισμός γονιδίων που επηρεάζονται από αυτά απαιτεί περαιτέρω έρευνα για την αποσαφήνιση των λειτουργικών τους ρόλων<sup>178,179</sup>. Παρακάτω ακολουθούν αντιπροσωπευτικά παραδείγματα γονιδίων που εμφανίζουν αλληλεπίδραση με CNVs στον καρκίνο του στομάχου (**Εικόνα 9**).

Το γονίδιο *PIK3CA* που εντοπίζεται στην περιοχή 3q26.3 του χρωμοσώματος 3 εμφανίζεται συχνά ενισχυμένο στον καρκίνο του στομάχου<sup>180</sup>. Η ενίσχυση αυτή φαίνεται να διαδραματίζει καθοριστικό ρόλο στην συνολική πρόγνωση της νόσου, ανεξαρτήτως του σταδίου του όγκου, υποδηλώνοντας την υψηλή της σημασία στην πολύπλοκη διαδικασία της καρκινογένεσης και καθιστώντας το συγκεκριμένο γονίδιο ως υψηλής προγνωστικής αξίας<sup>181,182</sup>. Επιπλέον, μελέτες έχουν κατατάξει την ενίσχυση του γονιδίου *S6K2* ως δείκτη κακής πρόγνωσης. Το συγκεκριμένο γονίδιο εμφανίζεται να απορυθμίζει τον μεταγραφικό παράγοντα mTOR στον καρκίνο του στομάχου<sup>183</sup>. Επίσης, ενίσχυση του γονιδίου *TNK2* στην περιοχή 3q29 αποτελεί πιθανό προγνωστικό δείκτη της ασθένειας, συμβάλλοντας στην αύξηση της κακοήθειας<sup>184</sup>. Το ογκοκατασταλτικό γονίδιο *APC* στην περιοχή 5q21-q22 έχει πρωταρχικό ρόλο σε ποικίλες κυτταρικές διεργασίες, όπως ο πολυμερισμός των μικροσωληνίσκων, η μεταγωγή σήματος και η κυτταρική προσκόλληση<sup>185</sup>. Το προϊόν του γονιδίου είναι αρνητικός ρυθμιστής του μονοπατιού WNT. Η απενεργοποίησή του προκαλεί συσσώρευση της β-κατενίνης και μεταγραφικής ενεργοποίησης γονιδίων σχετικών με τον κυτταρικό πολλαπλασιασμό. Ο γενετικός τύπος του *APC* εμφανίζει υψηλή συχνότητα ελλείψεων στον καρκίνο του στομάχου και ο μειωμένος αριθμός των αντιγράφων εμφανίζει υψηλή συσχέτιση με μετάσταση στους λεμφαδένες<sup>186</sup>. Πιο συγκεκριμένα, έλλειψη του γονιδίου είχε εντοπιστεί αρχικά σε γαστρικούς καρκίνους προχωρημένου σταδίου, γεγονός που το καθιστά σημαντικό στην εξέλιξη αλλά όχι στην έναρξη της νόσου. Ακόμη μια περίπτωση έλλειψης ογκοκατασταλτικού γονιδίου στον γαστρικό καρκίνο είναι αυτή του *IRX1* στην περιοχή 5p15.33. Το εν λόγω γονίδιο φυσιολογικά καταστέλλει διαδικασίες πολλαπλασιασμού, μετάστασης και ογκογένεσης. Η μείωση του αριθμού αντιγράφων του και κατά συνέπεια η μεταβολή στα επίπεδα έκφρασης του εμφανίζει συσχέτιση με μεθυλίωση υποκινητών<sup>57</sup>.

Το γονίδιο *MET* εντοπίζεται στην περιοχή 7q21 του χρωμοσώματος 7 η οποία κωδικοποιεί έναν υποδοχέα του αυξητικού παράγοντα των ηπατικών κυττάρων. Η δυσλειτουργία του μονοπατιού *MET* έχει παρατηρηθεί σε πολλούς τύπους καρκίνου, μεταξύ αυτών και ο καρκίνος του στομάχου. Η ενεργοποίηση της *MET* σηματοδότησης προκαλείται κυρίως από ενίσχυση του γονιδίου *MET*, προκαλώντας αύξηση του μεγέθους του όγκου, εισβολή στους παρακείμενους ιστούς και αγγειογένεση. Μάλιστα, ο αυξημένος αριθμός των αντιγράφων του θεωρείται πιο αποτελεσματικός προγνωστικός δείκτης από την αυξημένη ποσότητα της πρωτεΐνης για την ασθένεια<sup>187-189</sup>. Κατά συνέπεια, αποτελεί θεραπευτικό στόχο μοριακών αναστολέων. Γονιδιωματική ενίσχυση του γονιδίου *EGFR* σχετίζεται επίσης με υψηλό ρίσκο μετάστασης σε γαστρικούς όγκους και η ταυτοποίησή της συνδέεται συνήθως με μη ευνοϊκή πρόγνωση. Ενισχυμένη περιοχή του *EGFR* παρατηρείται σε ποσοστό άνω του 20% των περιπτώσεων γαστρικού καρκίνου, γεγονός που το καθιστά σημαντικό βιοδείκτη για την ασθένεια<sup>190-192</sup>.

Αύξηση του αριθμού των αντιγράφων του γονιδίου που κωδικοποιεί τον μεταγραφικό παράγοντα *c-MYC* φαίνεται να έχει επίδραση σε διαδικασίες πολλαπλασιασμού, διαφοροποίησης και απόπτωσης. Η απορρύθμιση του συγκεκριμένου γονιδίου είναι χαρακτηριστικό του μοριακού φαινοτύπου του γαστρικού καρκίνου, τόσο σε επίπεδο βιοψιών, όσο και κυτταρικών σειρών<sup>193,194</sup>. Αντίστοιχα με το *c-MYC*, το γονίδιο *POU5F1B* που βρίσκεται σε γειτονική περιοχή στο χρωμόσωμα 8, εμφανίζει επίσης αυξημένο αριθμό αντιγράφων. Η ενίσχυση του οδηγεί σε υπερέκφραση και ακόλουθη επαγωγή της καρκινογένεσης στο στομάχι<sup>195,196</sup>. Το γονίδιο *HER2* στην περιοχή 17q21 του χρωμοσώματος 17 εμφανίζει ενίσχυση που συσχετίζεται με αλληλεπίδραση με το *H. pylori* και προκαλεί καρκινογόνες αλλοιώσεις, σε συνδυασμό με περιβαλλοντικούς παράγοντες<sup>197-199</sup>. Το γονίδιο *TP53* στην περιοχή 17p13.1 κωδικοποιεί έναν κυρίαρχο ρυθμιστή γονιδιωματικής σταθερότητας<sup>200,201</sup>. Ως απάντηση στις βλάβες στο DNA, η πρωτεΐνη p53 πυροδοτεί έναν τεράστιο αριθμό κυτταρικών αποκρίσεων, συμπεριλαμβανομένων των διεργασιών επιδιόρθωσης του DNA, απόπτωσης, κυτταρικής διαφοροποίησης, μεταβολισμού, αγγειογένεσης και ανοσοαπόκρισης<sup>202-204</sup>. Σημαντική συσχέτιση μεταξύ ελλείψεων του γονιδίου *TP53* και εμφάνισης προκαρκινικών γαστρικών αλλοιώσεων, υποστηρίζει την εμπλοκή των CNVs στα πρώιμα στάδια της καρκινογένεσης<sup>205-208</sup>.



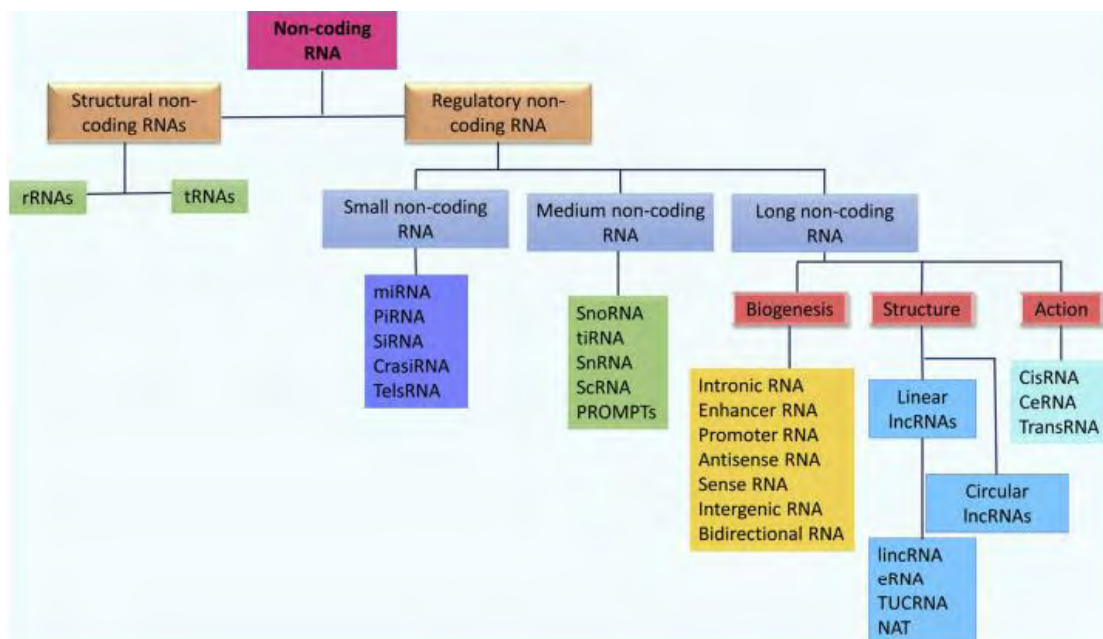
**Εικόνα 9:** Επισκόπηση των γονιδίων που εμφανίζουν CNVs σε συνδυασμό με τα μονοπάτια μεταγωγής σήματος, τα παθολογικά χαρακτηριστικά και τους θεραπευτικούς τους στόχους<sup>95</sup>.

Τη σχετική λίστα συμπληρώνουν και πολλές περιπτώσεις CNVs γενετικών τόπων που περιέχουν γονίδια, όπως ενισχύσεις των γονιδίων *AURKA*, *C20orf11*, *OCT1*, *FGFR2*, *CCND1* και *ORAOV1*, αλλά και έλλειψη του γονιδίου *GKN1*<sup>95</sup>. Πολύ σημαντική ενδέχεται να είναι και η συνεισφορά των CNVs που εντοπίζονται στο μιτοχονδριακό DNA (**mtDNA**). Πιο συγκεκριμένα, η μείωση στον αριθμό των αντιγράφων των γονιδίων είναι εμφανής στα κλινικοπαθολογικά στάδια III και IV του γαστρικού καρκίνου. Αρκετές δομικές ανακατατάξεις έχουν ταυτοποιηθεί στο mtDNA, όπως ελλείψεις μεγάλου αριθμού βάσεων και μεταλλάξεις στην περιοχή του βρόχου D (**D-loop region**)<sup>209</sup>. Η δυσλειτουργία των μιτοχονδρίων εξαιτίας της παρουσίας CNVs δύναται να παίζει καθοριστικό ρόλο στην κακοήγη έκβαση του γαστρικού καρκίνου. Αυτό εξηγείται από τη συμμετοχή των μιτοχονδριακών γονιδίων στην παραγωγή ενέργειας, στον μεταβολισμό και στον προγραμματισμένο κυτταρικό θάνατο<sup>210-212</sup>.

Οι αλλαγές στον αριθμό των αντιγράφων εντοπίζονται σε διαφορετικές χρωμοσωμικές περιοχές και επηρεάζουν την έκφραση ποικίλων γονιδίων, οδηγώντας τελικά στην ανάπτυξη γαστρικού καρκίνου<sup>95</sup>. Οι περισσότερες μελέτες εστιάζουν σε CNVs γονιδίων που κωδικοποιούν πρωτεΐνες. Παρόλα αυτά, τα τελευταία χρόνια η έρευνα έχει προσανατολιστεί σε αλλαγές στον αριθμό των αντιγράφων μη κωδικοποιητικών γονιδίων. Υπάρχουν ενδείξεις ότι τόσο τα μη κωδικοποιητικά RNA (**non-coding RNAs - ncRNAs**) όσο και τα μικρά RNA (**miRNAs**) μπορούν να επηρεάζονται όσον αφορά τη λειτουργία τους από την αλλαγή στον αριθμό των αντιγράφων τους στο γονιδίωμα<sup>213,214</sup>. Γι' αυτό και χρήζουν εκτενέστερης μελέτης με σκοπό την ανάδειξη τους ως χρήσιμους παράγοντες στην πρόοδο της νόσου.

### 1.3. Τα μακρά μη κωδικοποιητικά μόρια RNA (long-non coding RNAs)

Η μελέτη των λειτουργιών του RNA έχει ενταθεί τα τελευταία χρόνια με την ανακάλυψη ότι το γονιδίωμα των θηλαστικών υπόκειται σε εκτεταμένη μεταγραφή<sup>215</sup>. Αποτέλεσμα αυτού του φαινομένου είναι η έκφραση χιλιάδων μη κωδικοποιητικών μορίων RNA, τα οποία κατέχουν σημαντικούς ρόλους σε κυτταρικές διαδικασίες<sup>216,217</sup> (Εικόνα 10). Στην πραγματικότητα, τα γονίδια που κωδικοποιούν πρωτεΐνες αντιπροσωπεύουν ένα ελάχιστο ποσοστό του ανθρώπινου γονιδιώματος, περίπου το 2%, ενώ όλο το υπόλοιπο αντιστοιχεί σε μη κωδικοποιητικά RNA<sup>218</sup>. Η μελέτη των ncRNAs έχει εστιάσει κυρίως σε μια μικρή υποκατηγορία τους, στα miRNAs, τα οποία έχουν κύριο ρόλο στην ρύθμιση της γονιδιακής έκφρασης<sup>219</sup>. Τα τελευταία χρόνια η προσοχή έχει στραφεί στα μεγαλύτερα, τόσο σε μέγεθος, όσο και σε αριθμό, μακρά μη κωδικοποιητικά μόρια RNA (**long non-coding RNAs - lncRNAs**).



Εικόνα 10: Κατηγορίες των μη κωδικοποιητικών μορίων RNA<sup>220</sup>.

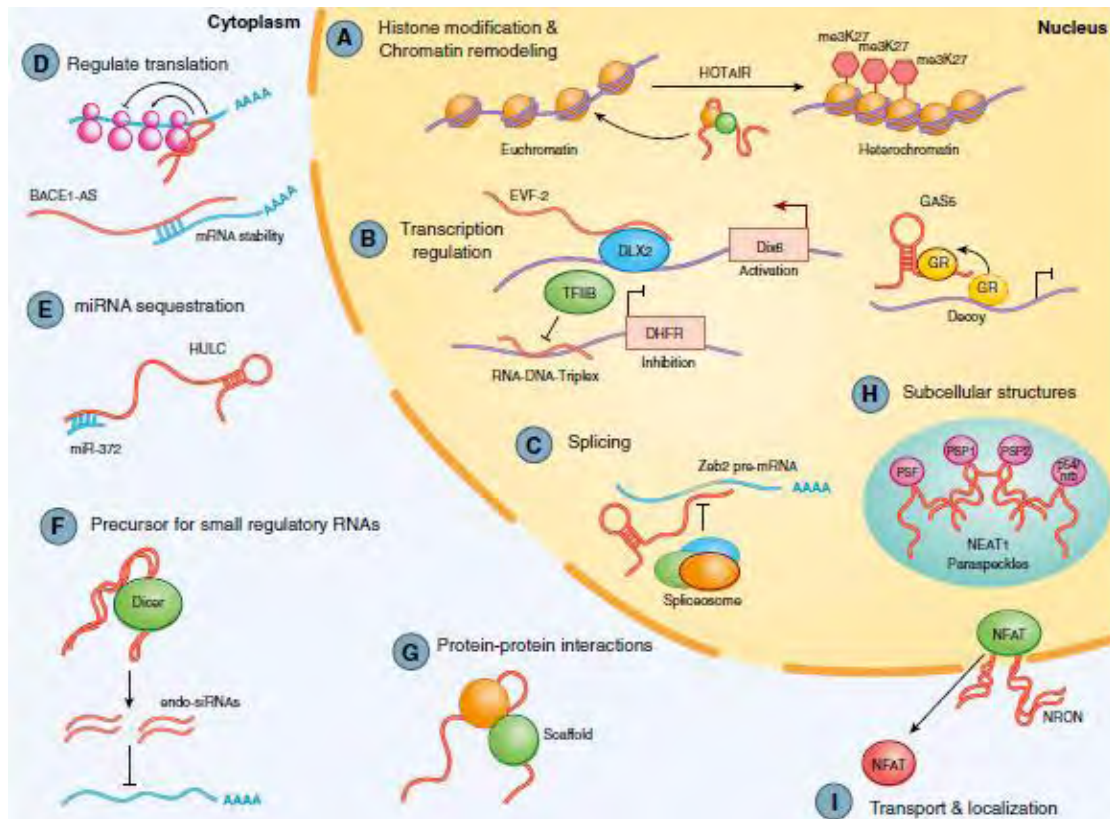
#### 1.3.1. Χαρακτηριστικά και κατηγοριοποίηση των lncRNAs

Τα lncRNAs συγκροτούν μια ετερογενή ομάδα μορίων RNA τα οποία ποικίλουν όσον αφορά στο μήκος (άνω των 200 νουκλεοτιδίων και άνω των 100kb), τον υποκυτταρικό εντοπισμό (στον πυρήνα, στο κυτταρόπλασμα ή και στα δύο), καθώς και τον λειτουργικό τους ρόλο σε ένα εύρος βιολογικών διεργασιών<sup>221</sup>. Η έκφραση τους περιορίζεται σε χαμηλότερα επίπεδα και εμφανίζει χαρακτηριστική ιστοειδικότητα και ειδοειδικότητα σε σχέση με τα κωδικοποιητικά μετάγραφα. Ομοίως με το mRNA, τα lncRNAs μεταγράφονται από την RNA πολυμεράση II (**RNA pol II**), διαθέτουν καλύπτρα στο 5' άκρο τους (**5' cap**) και πολυαδενυλιωμένο 3' άκρο (**poly-A 3' end**) και υφίστανται εναλλακτικό μάτισμα<sup>222,223</sup>. Επιπλέον μπορούν να προκύψουν και από τους δύο κλώνους του DNA, ανεξαρτήτως προσανατολισμού<sup>224</sup>. Περίπου το 80% του μεταγραφόμενου των θηλαστικών προβλέπεται ότι μεταγράφεται σε μη κωδικοποιητικά RNAs, ωστόσο ελάχιστος αριθμός αυτών έχει χαρακτηριστεί ως προς τη λειτουργία τους<sup>225</sup>. Σε αντίθεση με τα miRNAs, τα lncRNAs δεν φαίνεται να εμφανίζουν υψηλή συντήρηση εξελικτικά κι αυτός είναι ο λόγος που αρχικά χαρακτηρίστηκαν ως μεταγραφικός θόρυβος<sup>226,227</sup>. Παρόλα αυτά, συνεχώς αυξανόμενα

στοιχεία υποδεικνύουν ότι τα lncRNAs διαδραματίζουν καθοριστικό ρόλο σε μια μεγάλη ποικιλία βιολογικών διαδικασιών μεταξύ αυτών της γονιδιακής ρύθμισης, του πακεταρίσματος του γονιδιώματος, την οργάνωση της χρωματίνης, την αντιστάθμιση της γονιδιακής δόσης και του γονιδιωματικού εντυπώματος<sup>228-230</sup>. Επιτελούν κρίσιμους ρόλους στο έλεγχο της γονιδιακής έκφρασης κατά τη διάρκεια των διαδικασιών ανάπτυξης και διαφοροποίησης, ενώ ο συνολικός αριθμός των γενετικών τόπων τους στο γονιδίωμα αυξάνεται με την πολυπλοκότητα των οργανισμού, αποδεικνύοντας την σημασία των ρυθμιστικών μηχανισμών που βασίζονται στα lncRNAs, στην εξέλιξη των πολυκύτταρων οργανισμών<sup>231</sup>. Πιο συγκεκριμένα, τα lncRNAs μπορούν μεταξύ άλλων να κατηγοριοποιηθούν βάσει της θέσης τους στο γονιδίωμα ή στο εκάστοτε υποκυτταρικό διαμέρισμα<sup>232</sup>.

Η ενδογενής φύση των lncRNAs τους προσδίδει την διττή ικανότητα να λειτουργούν ως πρόσθετες πρωτεϊνών και να αξιοποιούν το φαινόμενο της συμπληρωματικότητας των βάσεων, για να συγκροτούν σύμπλοκα με στόχους ορισμένες περιοχές στο DNA ή μόρια RNA<sup>233-235</sup>. Σε αντίθεση με τα μικρά μη κωδικοποιητικά RNA (**small non-coding RNAs**) τα lncRNAs μπορούν να σχηματίσουν περίπλοκες δομές δευτεροταγούς ή υψηλότερης τάξης, αυξάνοντας την δυνατότητα για την αναγνώριση αλληλεπιδρώντων πρωτεϊνών και άλλων στόχων<sup>233-236</sup>. Πιο συγκεκριμένα, εντός του πυρήνα έχουν την ικανότητα να διευκολύνουν ή να παρεμποδίζουν την στρατολόγηση της RNA pol II, μεταγραφικών παραγόντων ή/και συμπαραγόντων σε υποκινητές γονιδίων, ελέγχοντας την μεταγραφή των γονιδίων στόχων. Επιπλέον, συμμετέχουν στη ρύθμιση του εναλλακτικού ματίσματος των πρώιμων mRNA<sup>237</sup>. Η ευέλικτη φύση της δευτεροταγούς τους δομής τους δίνει τη δυνατότητα να φέρουν σε επαφή πρωτεϊνικούς παράγοντες οι οποίοι ειδάλλως δεν θα μπορούσαν να αλληλεπιδράσουν<sup>233,234,238</sup>. Αυτή η λειτουργία τους ως κριώματα είναι επίσης σημαντική για την ενεργότητα και τον εντοπισμό της εκάστοτε πρωτεΐνης στόχου, καθώς και στην σταθεροποίηση υποκυτταρικών δομών<sup>239,240</sup>.

Τα lncRNAs μπορούν επίσης να οδηγήσουν στη συγκρότηση συμπλεγμάτων αναδιαμόρφωσης της χρωματίνης όπως το PRC2 (**Polycomb Repressive Complex 2**), στις κατάλληλες χρωματινικές περιοχές, παρέχοντας εξειδίκευση όσον αφορά το στόχο. Μέσω αυτών των αλληλεπιδράσεων, επιτυγχάνεται το πακετάρισμα της χρωματίνης σε τοπικό ή καθολικό επίπεδο, αλλά και η κατανομή των ευχρωματινικών και των ετεροχρωματινικών περιοχών<sup>241,242</sup>. Ακόμη, δύνανται να αλληλεπιδράσουν μέσω συμπληρωματικότητας με μόρια mRNA ή miRNA και να επηρεάσουν την σταθερότητα τους ή τη λειτουργία τους, διατηρώντας το αρχικό τους μέγεθος (**miRNA/mRNA sponges**) ή έπειτα από την μετατροπή τους σε siRNAs (**small interfering RNAs**)<sup>243-247</sup>. Τέλος, κάποια lncRNAs μπορούν να κωδικοποιούν για μικρά πεπτίδια και πρωτεΐνες μικρού μεγέθους<sup>248,249</sup> (**Εικόνα 11**).



Εικόνα 11: Σχηματική απεικόνιση των κυτταρικών λειτουργιών των lncRNAs<sup>221</sup>.

### Κατηγοριοποίηση των lncRNAs ανάλογα με τη θέση τους στο γονιδίωμα

Ανάλογα με την οργάνωσή τους και τη σχετική τους θέση σε σχέση με τα γονίδια που κωδικοποιούν πρωτεΐνες, τα lncRNAs κατηγοριοποιούνται στις εξής γενικές κατηγορίες<sup>250</sup> (Εικόνα 12):

**Ιντρονικά lncRNAs (Intronic lncRNAs):** Εντοπίζονται στο εσωτερικό ενός ιντρονίου μιας κωδικής περιοχής.

**Μεσογονιδιακά lncRNAs (Intergenic lncRNAs - lincRNAs):** Παρεμβάλλονται μεταξύ των αλληλουχιών δύο κωδικοποιητικών γονιδίων.

**Νοηματικά lncRNAs (Sense lncRNAs):** Μεταγράφονται από τον μη κωδικό ή μεταγραφόμενο κλώνο σε περιοχές γονιδίων που κωδικοποιούν πρωτεΐνες. Ενδέχεται να αλληλοεπικαλύπτονται με περιοχές εξωνίων ή ιντρονίων.

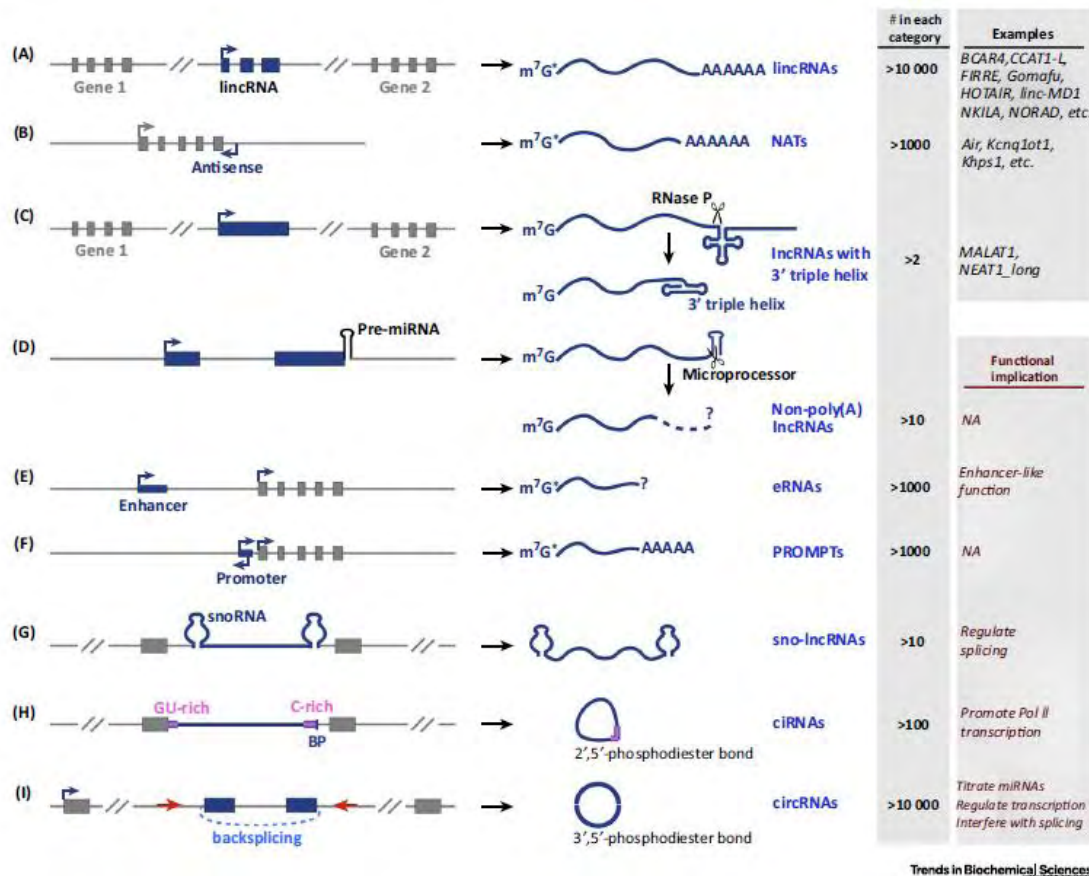
**Αντινοηματικά lncRNAs (Antisense lncRNAs):** Μεταγράφονται από τον κωδικό ή μη μεταγραφόμενο κλώνο σε περιοχές γονιδίων που κωδικοποιούν πρωτεΐνες. Ενδέχεται να εμφανίζουν συμπληρωματικότητα με περιοχές εξονίων ή ιντρονίων γονιδίων που κωδικοποιούν πρωτεΐνες και μεταγράφονται από την συμπληρωματική αλυσίδα του DNA.

**Αμφίδρομα lncRNAs (Bidirectional lncRNAs):** Εντοπίζονται σε περιοχές 1kb ανοδικά ή καθοδικά υποκινητών, έχοντας την ίδια κατεύθυνση (**convergent**) ή αντίθετη κατεύθυνση (**divergent**) με ένα πλησίον κωδικοποιητικό γονίδιο.

**lncRNAs ενισχυτών (Enhancer lncRNAs):** Συνεντοπίζονται και μεταγράφονται από αλληλουχίες ενισχυτών στο γονιδίωμα.

**Επικαλύπτοντα lncRNAs (Overlapping lncRNAs):** Είναι παρόντα εντός των ορίων γονιδίων που κωδικοποιούν πρωτεΐνες. Κάποια από αυτά ενδέχεται να αποτελούν παραλλαγές των mRNAs που παράγονται από αυτά τα γονίδια, αλλά να μην διαθέτουν λειτουργικό πλαίσιο ανάγνωσης (**open reading frame - ORF**).

**Κυκλικά lncRNAs (Circular lncRNAs - circRNAs):** Προκύπτουν συνήθως από σφάλματα κατά τη διαδικασία του εναλλακτικού ματίσματος. Μπορούν να περιέχουν ένα ή περισσότερα εξώνια αλλά και ιντρονικές περιοχές.



**Εικόνα 12:** Κατηγορίες των lncRNAs ανάλογα με την οργάνωσή τους στο γονιδίωμα<sup>251</sup>.

### Κατηγοριοποίηση των lncRNAs ανάλογα με τον υποκυτταρικό τους εντοπισμό

Όπως οι πρωτεΐνες, έτσι και τα lncRNAs βασίζουν την λειτουργία τους ανάλογα με το υποκυτταρικό διαμέρισμα στο οποίο βρίσκονται<sup>251</sup>. Πολλά lncRNAs αναγνωρίζονται ως σημαντικοί συντελεστές πυρηνικών διεργασιών και παρουσιάζουν εξειδικευμένο υποκυτταρικό εντοπισμό<sup>235,252</sup>. Άλλα lncRNAs προϋποθέτουν την έξοδό τους στο κυτταρόπλασμα προκειμένου να επιτελέσουν τον ρυθμιστικό τους ρόλο<sup>253-255</sup>. Τα lncRNAs υποκατηγοριοποιούνται περαιτέρω σε τρεις κύριες κατηγορίες ανάλογα με την απόσταση του γενετικού τους τόπου από το γονίδιο-στόχο τους: στα αμιγώς πυρηνικά που δρουν κοντά στον γενετικό τους τόπο (**cis-acting lncRNAs**), στα αμιγώς πυρηνικά που δρουν μακριά από τον γενετικό τους τόπο (**trans-acting lncRNAs**) και στα κυτταροπλασματικά **trans-acting lncRNAs**<sup>251,256</sup>. Αξίζει να σημειωθεί ότι οι παραπάνω ρόλοι δεν αντιπροσωπεύουν το σύνολο των περιπτώσεων, καθώς υπάρχουν περιπτώσεις lncRNAs που δρουν τόσο στον πυρήνα όσο και στο κυτταρόπλασμα<sup>257,258</sup>.

## Πυρηνικά lncRNAs που δρουν κοντά στον γενετικό τους τόπο (*cis*-acting lncRNAs)

Τα συγκεκριμένα lncRNAs επιτελούν τους ρυθμιστικούς τους ρόλους με μια ποικιλία μηχανισμών<sup>259</sup>. Για παράδειγμα, μπορούν να πυροδοτήσουν την γονιδιακή έκφραση στρατολογώντας μεταγραφικούς παράγοντες ή μόρια αναδιαμόρφωσης της χρωματίνης στα γειτονικά τους γονίδια. Σχηματίζουν δομές τριπλής έλικας DNA-RNA, που συγκρατούν το ίδιο το lncRNA και τις υπόλοιπες πρωτεΐνες-τελεστές στην περιοχή στόχο, η οποία μπορεί να είναι ένας υποκινητής<sup>260</sup>. Επιπλέον, ρυθμιστικά RNA που μεταγράφονται από γειτονικές περιοχές ενισχυτών, παραμένουν προσδεδεμένα στη χρωματίνη και συμβάλλουν στην σταθεροποίηση της πρόσδεσης μεταγραφικών παραγόντων<sup>261</sup>. Ειδικότερα, το lncRNA CCAT1-L μεταγράφεται από έναν υπερενισχυτή (**super-enhancer**) ανοδικά του γονιδίου *MYC*. Το CCAT1-L δρα *in cis*, αλληλεπιδρά με τον παράγοντα αναδιαμόρφωσης της χρωματίνης CTCF, και προάγει τον σχηματισμό βρόχων στη χρωματίνη. Με αυτόν τον τρόπο, ο υποκινητής του *MYC* και ο υπερενισχυτής έρχονται σε κοντινή απόσταση και αλληλεπιδρούν, οδηγώντας σε αυξημένη έκφραση του γονιδίου *MYC* στον καρκίνο του παχέος εντέρου<sup>262,263</sup> (**Εικόνα 13A**).

Πολλά καλά χαρακτηρισμένα φυσικά αντισηματικά μετάγραφα (**Natural Antisense Transcripts - NATs**) προσεγγίζουν μεταγραφικά ενεργές περιοχές και επηρεάζουν την έκφραση γειτονικών γονιδίων<sup>264-266</sup>. Όσον αφορά το εντύπωμα των θηλαστικών, τα lncRNAs *Air* και *Kcnq1ot1* εμπλέκονται στην σίγηση γειτονικών τους γονιδίων. Το lncRNA *Air* στρατολογεί την μεθυλοτρανσφεράση G9a της λυσίνης 9 της ιστόνης 3 (H3K9) προκειμένου να διατηρήσει την σίγηση ενός συνόλου γονιδίων, μεταξύ αυτών και του γενετικού τόπου *Igf2r*<sup>266</sup>. Αντίστοιχα, το *Kcnq1ot1* στρατολογεί την ίδια μεθυλοτρανσφεράση και το σύμπλοκο PRC2 για τη σίγηση γονιδίων στον γενετικό τόπο *Kcnq1*<sup>264,265</sup>. Τέλος, το lncRNA *Khps1* αυξάνει την έκφραση του γονιδίου *SPHK1* σχηματίζοντας δομή τριπλής έλικας και οδηγώντας σε αλλαγές στην χρωμοσωμική δομή<sup>267</sup>.

Μερικά από τα *cis*-acting lncRNAs μπορούν να επηρεάσουν την γονιδιακή έκφραση και *in trans*. Το lncRNA *FIRRE* μεταγράφεται από το χρωμόσωμα X και συμμετέχει στην διαμόρφωση αλληλεπιδράσεων μεταξύ χρωμοσωμάτων σε μεταγραφικά ενεργές περιοχές<sup>268</sup>. Αυτός ο χωρικός συνεντοπισμός επιτυγχάνεται μέσω της τρισδιάστατης οργάνωσης των χρωμοσωμάτων, η οποία επιτρέπει αλληλεπιδράσεις μεγάλης εμβέλειας μεταξύ ρυθμιστικών στοιχείων και γονιδίων στόχων<sup>268</sup>. Επιπλέον, διακρίνεται ακόμη ένας μηχανισμός μέσω του οποίου τα *cis*-acting lncRNAs μπορούν να λειτουργήσουν ως δολώματα (**decoys**) για πρωτεΐνες που προσδένονται στο RNA (**RNA-binding Proteins - RBPs**), ώστε να εξασφαλίσουν την γονιδιακή ρύθμιση *in trans*. Παράδειγμα αποτελεί μια κατηγορία μικρών lncRNAs του πυρηνοειδούς (**small nucleolar RNA-ended lncRNAs - sno-lncRNAs**) σε ανθρώπινα εμβρυικά βλαστικά κύτταρα (**human embryonic stem cells - hESC**). Τα μόρια αυτά εκφράζονται σε μεγάλες ποσότητες και συσσωρεύονται στην περιοχή της σύνθεσής τους, συμμετέχοντας σε διαδικασίες εναλλακτικού ματίσματος<sup>269,270</sup> (**Εικόνα 13B**).

Ο μηχανισμός με τον οποίο τα lncRNAs περιορίζονται στην περιοχή του γονιδιακού τους τόπου είναι μέχρι στιγμής άγνωστος, ωστόσο συχνά τέτοια lncRNAs παράγουν μικρό αριθμό μεταγράφων ανά κύτταρο συνεπώς η δράση τους εκ των πραγμάτων περιορίζεται στο γειτονικό τους γονίδιο. Η πρόσδεση σε κοντινούς παράγοντες<sup>262,264,265,271,272</sup> και ο σχηματισμός δομών τριπλής έλικας RNA-DNA<sup>267</sup> συνιστούν δύο επιπρόσθετους μηχανισμούς συσώρευσής τους σε τοπικό επίπεδο. Παρόλα αυτά, εξετάζεται το γεγονός ο περιορισμός τους να συνδέεται και με αλληλεπιδράσεις με την RNA pol II, κατά τη διάρκεια της μεταγραφής<sup>272</sup>.



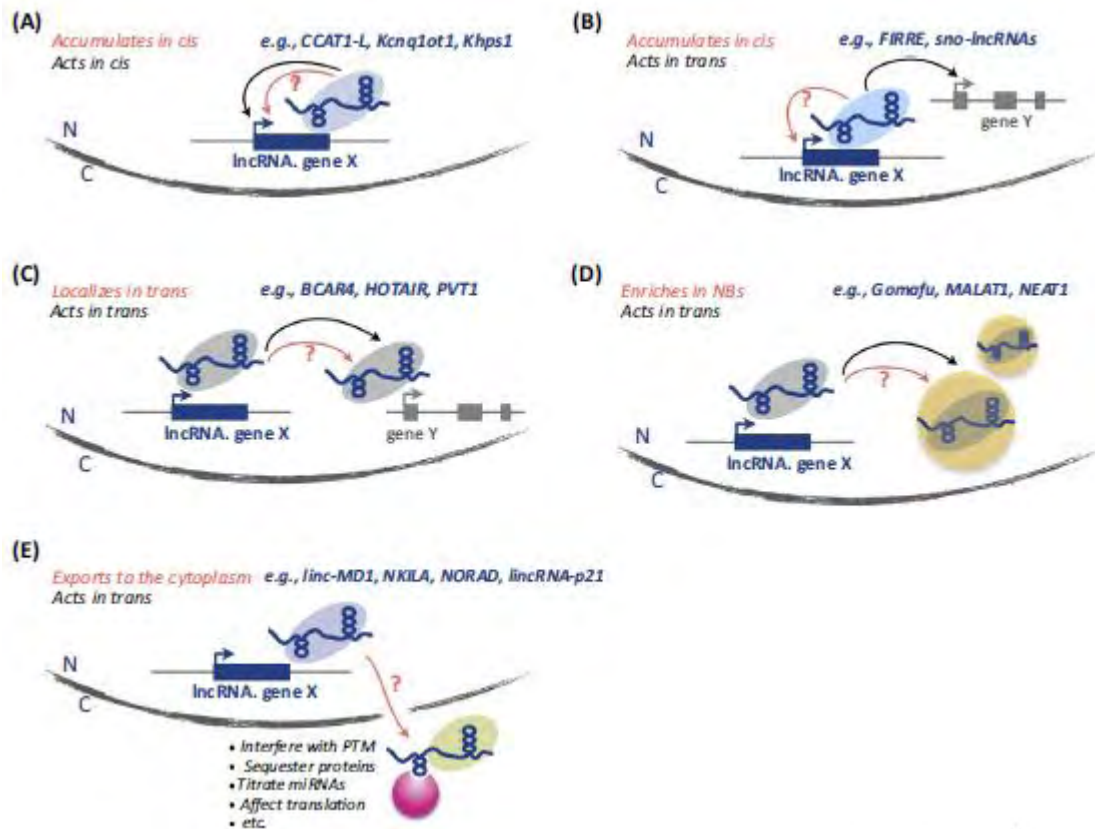
## Πυρηνικά lncRNAs που δρουν ανεξάρτητα από τον γενετικό τους τόπο (*trans-acting* lncRNAs)

Η εν λόγω κατηγορία lncRNAs περιλαμβάνει μόρια που εντοπίζονται επίσης στον πυρήνα, ωστόσο η αποτελεσματική ρυθμιστική τους δράση προϋποθέτει την μετακίνησή από τις θέσεις σύνθεσής τους<sup>273,274</sup>. Το lncRNA HOTAIR (**HOX transcript antisense RNA**) είναι ένα από τα πιο καλά μελετημένα μόρια αυτής της κατηγορίας<sup>275-277</sup>. Μεταγράφεται από τον γενετικό τόπο *HOXC*, ωστόσο η δράση του ανιχνεύεται στον γενετικό τόπο *HOXD*, στον οποίο καταστέλλει την διαδικασία της μεταγραφής *in trans*, μέσω της στρατολόγησης του συμπλόκου PRC2 κατά την εγκαθίδρυση του αναπτυξιακού προτύπου<sup>238,242</sup>. Η έκφραση του HOTAIR επηρεάζει το γονιδίωμα και σε επιγενετικό επίπεδο, οδηγώντας σε τριμεθυλιώσεις της λυσίνης 27 της ιστόνης 3 (H3K27me3), χαρακτηριστικής ένδειξης καταστολής της έκφρασης<sup>278</sup>. Επιπλέον, το HOTAIR προσδένεται σε ένα μεγάλο εύρος περιοχών στο γονιδίωμα, μέσω μιας περιοχής πλούσιας σε επαναλήψεις GA<sup>279</sup>. Οι επαναλήψεις αυτές πιστεύεται ότι συμβάλλουν στον σχηματισμό τριπλής έλικας, η οποία αλληλεπιδρά με το σύμπλοκο αναδιαμόρφωσης της χρωματίνης. Τα lncRNAs BCAR4 και PVT1 αποτελούν ακόμη δύο πυρηνικά lncRNAs, τα οποία εντοπίζονται ως επί το πλείστον στο πυρηνόπλασμα. Το BCAR4 αλληλεπιδρά με τους μεταγραφικούς παράγοντες SNIP1 και PNUITS μέσω δύο ξεχωριστών περιοχών στο γονιδίωμα, επηρεάζοντας την επιγενετική ρύθμιση ως απόκριση στη σηματοδότηση από χημειοκίνες στον καρκίνο του μαστού<sup>280</sup>. Το PVT1 αλληλεπιδρά με τον μεταγραφικό παράγοντα MYC στον πυρήνα, αυξάνοντας την σταθερότητα του και κατά συνέπεια τα επίπεδά του σε ένα μεγάλο εύρος καρκίνων<sup>281</sup> (**Εικόνα 13C**).

Εκτός από τις παραπάνω περιπτώσεις μορίων οι οποίες εμφανίζουν καθολική παρουσία στον πυρήνα, διακρίνονται και lncRNAs τα οποία συσσωρεύονται στοχευμένα σε συγκεκριμένα νουκλεοσώματα, συμμετέχοντας σε ριβοπρωτεϊνικά σύμπλοκα, τα οποία συντονίζουν πυρηνικές διαδικασίες<sup>282</sup>. Για παράδειγμα, το MALAT (**Metastasis-Associated Lung Adenocarcinoma Transcript 1**) είναι ένα lncRNA που εντοπίζεται σε περιοχές πλούσιες σε παράγοντες ματίσματος (**nuclear speckles**), και έχει αποδειχθεί ότι συμμετέχει σε ογκογόνες διαδικασίες<sup>283-285</sup>. Το NEAT1 είναι επίσης ένα lncRNA που εντοπίζεται κοντά σε ριβονουκλεοπρωτεϊνικά σύμπλοκα (**long noncoding ribonucleoprotein (lncRNP) complexes**), εξασφαλίζοντας την ακεραιότητά τους<sup>240,286</sup>. Πιο συγκεκριμένα, η μεταγραφή τους αποτελεί έναυσμα για τη συναρμολόγηση των σωματιδίων<sup>287-290</sup>. Με αυτό τον τρόπο, ένζυμα όπως η RNA pol II συμβάλλουν κατά ένα σημαντικό ποσοστό στην ελεγχόμενη κατανομή των lncRNAs στον πυρήνα μέσω ποικίλων μηχανισμών (**Εικόνα 13D**).

## Κυτταροπλασματικά lncRNAs

Η τρίτη και τελευταία κατηγορία των lncRNAs είναι αυτή που η δράση τους απαιτεί την έξοδό τους στο κυτταρόπλασμα. Τα κυτταροπλασματικά lncRNAs μπορούν να επηρεάσουν την γονιδιακή έκφραση με πολλούς τρόπους<sup>253</sup>. Καταρχάς, αλληλεπιδρώντας με πρωτεΐνες σε μεταφραστικό αλλά και σε μετα-μεταφραστικό επίπεδο, με αποτέλεσμα την ενίσχυση της μεταγωγής σήματος<sup>291,292</sup>. Επιπλέον η δράση τους περιλαμβάνει τους ρόλους τους ως δολώματα τόσο ενάντια σε mRNAs, όσο και σε πρωτεΐνες, προκειμένου να περιορίσουν την δράση τους<sup>293-295</sup>. Τέλος, η δράση των κυτταροπλασματικών lncRNAs μπορεί να έγκειται και στην παραγωγή μικρών πεπτιδίων ή πρωτεϊνών, οι οποίες μπορούν να αξιοποιηθούν ως πεπτίδια οδηγοί ή αλλωστερικοί ανταγωνιστές ενζύμων<sup>296-300</sup> (**Εικόνα 13E**).



Εικόνα 13: Λειτουργίες των lncRNAs σε σχέση με τον υποκυτταρικό τους εντοπισμό<sup>251</sup>.

### Ο ρόλος των lncRNAs στον καρκίνο του στομάχου

Το φαινόμενο της μη ελεγχόμενης γονιδιακής έκφρασης αποτελεί ένα από τα θεμέλια γεγονότα της καρκινογένεσης<sup>223</sup>. Ενδελεχείς μελέτες πάνω στα αίτια της πρόκλησης του καρκίνου έχουν αναδείξει την παραλλακτικότητα μη κωδικών περιοχών ως υπεύθυνη για την ευαισθησία σε ένα μεγάλο αριθμό ογκογόνων παραγόντων<sup>301</sup>. Ένας μεγάλος αριθμός γενετικών τόπων που φιλοξενούν μη κωδικοποιητικά στοιχεία, και πιο συγκεκριμένα lncRNAs, δείχνουν να κατέχουν ζωτικούς ρόλους στην έναρξη και την εξέλιξη της ογκογένεσης<sup>302</sup>. Ως γνωστόν, τα lncRNAs μπορούν να αποτελέσουν μέρος πολλών βιολογικών διεργασιών εντός του κυττάρου, οι οποίες αφορούν ορόσημα του καρκίνου όπως η κυτταρική διαίρεση και η διαφοροποίηση<sup>223</sup>. Πολλά από αυτά τα μόρια έχουν εντοπιστεί να εμπλέκονται και στη γαστρική καρκινογένεση, είτε ως ογκοεπαγωγικά, είτε ως ογκοκατασταλτικά γονίδια<sup>303</sup>. Για παράδειγμα, υψηλά επίπεδα του lncRNA HOTAIR σχετίζονται με μετάσταση στους λεμφαδένες, εισβολή στα αγγεία και μείωση του προσδόκιμου ζωής στον γαστρικό καρκίνο<sup>304</sup>. Η εμφανής σύνδεση του με την διαδικασία της μετάστασης τον καθιστά έναν βασικό συμμετέχοντα στο μηχανισμό μετατροπής των επιθηλιακών κυττάρων σε μεσεγχυματικά (**Epithelial to Mesenchymal Transition - EMT**). Η καταστολή του αποτελεί έναν πιθανό τρόπο αναστροφής αυτής της διαδικασίας<sup>305</sup>. Επιπλέον, το lncRNA H19, εμφανίζει υψηλή συσχέτιση με διαδικασίες πολλαπλασιασμού, μετανάστευσης και εισβολής κυττάρων σε παρακείμενους ιστούς μέσω της υπερέκφρασής του στην νόσο. Σημαντικό χαρακτηριστικό του H19 είναι ο εντοπισμός του στο πλάσμα των ασθενών, καθιστώντας το εύκολα ανιχνεύσιμο<sup>306</sup>. Μείωση στην έκφραση του GAS5 (**Growth arrest-specific transcript 5**) εμφανίζει θετική συσχέτιση με το μέγεθος και το κλινικό στάδιο του όγκου, καθώς και με την εισβολή σε εγγύς λεμφαδένες. Ωστόσο, η αύξηση στην έκφραση

του κατά την χρήση ειδικής θεραπευτικής αγωγής φαίνεται να έχει θετικά αποτελέσματα στην αντιμετώπιση της νόσου<sup>307</sup>. Το lncRNA MEG3 (**Maternally Expressed Gene 3**), εμφανίζει επίσης μειωμένη έκφραση στον γαστρικό καρκίνο συγκριτικά με παρακείμενους φυσιολογικούς ιστούς και αποτελεί δείκτη κακής πρόγνωσης της ασθένειας<sup>308</sup>. Η δράση του έγκειται στην αλληλεπίδρασή του με miRNAs, καθώς και με πρωτεΐνες όπως η p53 για την ρύθμιση της κακοήθους δράσης του<sup>309</sup>. Επιπλέον παραδείγματα αποτελούν τα LINC00152 ANRIL και MALAT1 των οποίων η υπερέκφραση συνδέεται στενά με την επαγωγή της μετάστασης και την παρεμπόδιση της δράσης ογκοκατασταλτικών γονιδίων<sup>310</sup>.

Καθένα από τα παραπάνω παραδείγματα αναδεικνύει τον καθοριστικό ρόλο των lncRNAs στην γαστρική καρκινογένεση. Τόσο το ιστοειδικό πρότυπο δράσης τους, όσο και το μεγάλο εύρος των διαδικασιών στις οποίες εμπλέκονται τους καθιστά εξαιρετικούς διαγνωστικούς και προγνωστικούς βιοδείκτες, καθώς και θεραπευτικούς στόχους για τον καρκίνο του στομάχου<sup>230,311</sup>. Παράλληλα, διακρίνονται πολλοί γενετικοί τόποι που ενισχύονται στη νόσο. Επομένως, η συνδυαστική τους μελέτη στην επαγωγή της νόσου σε διαγνωστικά εργαλεία και σε νέες θεραπείες καθίσταται ιδιαίτερα χρήσιμη.

## 2. ΕΡΕΥΝΗΤΙΚΗ ΥΠΟΘΕΣΗ ΚΑΙ ΣΚΟΠΟΣ

## Ερευνητική υπόθεση

Τα CNVs φαίνεται να διαδραματίζουν καθοριστικό ρόλο στην πυροδότηση ογκογόνων διαδικασιών στον γαστρικό καρκίνο, τροποποιώντας την έκφραση μιας πληθώρας καλά χαρακτηρισμένων γονιδίων. Η μεταβολή του αριθμού των αντιγράφων μπορεί να οδηγήσει δυνητικά τόσο στην αύξηση όσο και στη μείωση των επιπέδων έκφρασης. Μέχρι και σήμερα, η πλειοψηφία των καταγεγραμμένων περιπτώσεων που συσχετίζονται με CNVs αφορά γονίδια που κωδικοποιούν πρωτεΐνες. Ελάχιστες περιπτώσεις καρκίνων υποδεικνύουν οποιαδήποτε εμπλοκή τους με τη δράση μη κωδικοποιητικών γονιδίων και ειδικότερα με lncRNAs. Σκοπός της παρούσας πτυχιακής είναι η ταυτοποίηση και ο χαρακτηρισμός ανάλογων lncRNA μορίων, τα οποία παρουσιάζουν ενίσχυση του γενετικού τους τύπου στον καρκίνο του στομάχου. Η μελέτη τους είναι εύλογη, καθώς παρουσιάζουν ένα ιστοειδικό πρότυπο δράσης σε ένα ευρύ φάσμα ρυθμιστικών διεργασιών, υπό το πρίσμα του συντονισμού και της οργάνωσης τους, τόσο σε επίπεδο φυσιολογικών, όσο και καρκινικών ιστών. Κατά συνέπεια, το ενδιαφέρον έγκειται στην ανάπτυξη νέων αξιόπιστων προγνωστικών και διαγνωστικών βιοδεικτών καθώς και καινοτόμων θεραπευτικών στόχων. Η υπόθεση στηρίζεται σε βιοπληροφορική ανάλυση δεδομένων έκφρασης lncRNA μορίων που δεν έχουν χαρακτηριστεί στο παρελθόν.

## Πειραματικός σκοπός

Το πειραματικό σκέλος της παρούσας πτυχιακής είναι αφιερωμένο στην ανάπτυξη μοριακών εργαλείων για την καταστολή της έκφρασης lncRNA μορίων στόχων σε καρκινικές σειρές στομάχου, με σκοπό την μελέτη και την αποσαφήνιση του ρόλου τους στον γαστρικό καρκίνο. Πιο συγκεκριμένα, περιλαμβάνει τον σχεδιασμό μορίων shRNA που στοχεύουν στα lncRNA στόχους, και την κλωνοποίησή τους στον φορέα κλωνοποίησης pSiCORmod-PGK-puro. Ο εν λόγω φορέας αξιοποιείται για τον μετασηματισμό καρκινικών κυττάρων με στόχο την παροδική καταστολή γονιδίων-στόχων. Επιπλέον διεξάγονται πειράματα μεταγραφικής σταθερότητας και υποκυτταρικής κλασμάτωσης, με σκοπό τον όσο το δυνατό πιο ολοκληρωμένο χαρακτηρισμό των lncRNA στόχων. Ο έλεγχος του φαινοτύπου θα πραγματοποιηθεί έπειτα από την πραγματοποίηση επιτυχημένης καταστολής μέσω των shRNAs, για πιθανό εντοπισμό πιθανών μεταβολών στη μορφολογία, αλλά και ως προς το πολλαπλασιαστικό δυναμικό. Μελλοντικό στόχο αποτελεί επίσης η αύξηση των επιπέδων της έκφρασης των lncRNA στόχων σε περισσότερες κυτταρικές σειρές, στις οποίες η έκφραση τους είναι μηδενική, για τον πλήρη καθορισμό των ιδιοτήτων τους και την ευκολότερη δυνατή μελέτη τους.

## 3. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ



### 3.1.3. Κυτταρικές σειρές

Οι κυτταρικές σειρές που χρησιμοποιήθηκαν κατά την διάρκεια των πειραμάτων ήταν οι εξής:

Κυτταρική σειρά	Οργανισμός	Ασθένεια	Ιστοπαθολογία	Πρόέλευση
HFE	Άνθρωπος	-	Φυσιολογικό επιθήλιο στομάχου	Φυσιολογικό επιθήλιο στομάχου
KATO III	Άνθρωπος	Καρκίνος στομάχου	Αδενοκαρκίνωμα σφραγιστήρος δακτυλίου	Πλευρική μετάσταση
MKN45	Άνθρωπος	Καρκίνος στομάχου	Φτωχά συνεκτικό αδενοκαρκίνωμα	Εστία μετάστασης: Ήπαρ
NUGC4	Άνθρωπος	Καρκίνος στομάχου	Φτωχά συνεκτικό αδενοκαρκίνωμα	Εστία μετάστασης: Παραγαστρικοί λεμφαδένες
HT29	Άνθρωπος	Καρκίνος παχέος εντέρου	Μερικώς συνεκτικό αδενοκαρκίνωμα	Κόλον
HCT116	Άνθρωπος	Καρκίνος παχέος εντέρου	-	Εστία μετάστασης: Παραγαστρικοί λεμφαδένες
MCF7	Άνθρωπος	Καρκίνος μαστού	Φτωχά συνεκτικό αδενοκαρκίνωμα	Πλευρική μετάσταση: Μαστός
SKBR3	Άνθρωπος	Καρκίνος μαστού	Φτωχά συνεκτικό αδενοκαρκίνωμα	Πλευρική μετάσταση: Μαστικοί αδένες

## 3.2. Μέθοδοι

### 3.2.1. Βιοπληροφορική ανάλυση

Στα πλαίσια της επιλογής του κατάλληλου μεταγράφου προς χαρακτηρισμό, πραγματοποιήθηκε ανάλυση δεδομένων RNA-seq, τόσο από φυσιολογικές και καρκινικές σειρές στομάχου, όσο και από δείγματα βιοψιών φυσιολογικού και καρκινικού γαστρικού ιστού. Στόχος της βιοπληροφορικής ανάλυσης ήταν να οπτικοποιηθεί η έκφραση του CENTRIC2 σε ένα ευρύ σύνολο δειγμάτων. Για τον υπολογισμό και την επιβεβαίωση του κωδικοποιητικού δυναμικού του CENTRIC2 χρησιμοποιήθηκαν δύο προγράμματα, το CPC2 (<http://cpc2.cbi.pku.edu.cn/>) και το CPAT (<http://lilab.research.bcm.edu/cpat/>). Τέλος, η ανάλυση των δεδομένων των q-PCR που διεξήχθησαν πραγματοποιήθηκε μέσω του λογισμικού CFX Manager της Bio-Rad.

### 3.2.2. Μετασηματισμός βακτηριακών στελεχών και απομόνωση πλασμιδιακού DNA

Το βακτηριακό στέλεχος που επιλέχθηκε προς μετασηματισμό ήταν το στέλεχος XL1 Blue του βακτηρίου E.coli. Το συγκεκριμένο στέλεχος δίνει τη δυνατότητα απομόνωσης ακέραιου πλασμιδιακού DNA, καθώς φέρει μεταλλάξεις οι οποίες αποτρέπουν τα γεγονότα ανασυνδυασμού. Επιπλέον, διαθέτει μεταλλάξεις που επιτρέπουν τον μετασηματισμό πλασμιδίων συμβατών με σύστημα διαμόλυνσης με λεντιούς, σημειώνοντας υψηλά ποσοστά επιτυχίας. Τα βακτηριακά κύτταρα λαμβάνονται από θερμοκρασία -80°C και



τοποθετούνται στον πάγο (-4°C) για 20 λεπτά, ώστε να προσαρμοστούν ήπια σε θερμοκρασία δωματίου και να αναμειχθούν με τον πλασμιδιακό φορέα. Στη συνέχεια, το μείγμα πλασμιδίου-βακτηρίων τοποθετείται σε υδατόλουτρο (42 °C) για 50 δευτερόλεπτα. Σε αυτό το χρονικό διάστημα τα βακτήρια υφίστανται θερμικό σοκ (heat shock), το οποίο διαταράσσει τις κυτταρικές του μεμβράνες, δημιουργώντας πόρους. Κατά συνέπεια, η είσοδος του πλασμιδίου στα βακτηριακά κύτταρα καθίσταται δυνατή. Ακολουθεί μεταφορά στον πάγο για 2 λεπτά και στη συνέχεια σε 900 μl θρεπτικό SOB απουσία αντιβιοτικού. Έπειτα, τοποθετούνται για 50 λεπτά σε επωαστήρα στους 37 °C. Σκοπός της επώασης είναι ο διπλασιασμός των βακτηρίων τουλάχιστον μια φορά (ο φυσιολογικός χρόνος διπλασιασμού του E.coli είναι περίπου 40 λεπτά), καθώς και η μεταγραφή του συνόλου των γονιδίων τους, συμπεριλαμβανομένου και του γονιδίου ανθεκτικότητας σε αντιβιοτικό. Με την περάτωση της επώασης, πραγματοποιείται επίστρωση των βακτηρίων σε τρυβλία με LB agar παρουσία αμπικιλίνης με στόχο την επιλογή αποικιών που έχουν λάβει επιτυχώς το ανασυνδυασμένο πλασμίδιο. Το τρυβλίο τοποθετείται στον επωαστήρα για ολονύχτια επώαση στους 37 °C. Την επόμενη μέρα, απομονώνονται μοναδιαίες αποικίες, οι οποίες τοποθετούνται σε 3ml θρεπτικού LB παρουσία κατάλληλου αντιβιοτικού και εισάγονται σε επωαστήρα υπό ανάδευση στους 37 °C. Μετά από διάστημα 8 ωρών και ανάπτυξης μιας κρίσιμης μάζας κυττάρων, πραγματοποιείται μεταφορά ποσότητας σε ογκομετρική φιάλη 120 ml LB παρουσία αντιβιοτικού. Η καλλιέργεια τοποθετείται σε αναδευόμενο επωαστήρα στους 37°C για ολονύχτια επώαση. Με την ολοκλήρωση της επώασης, τα βακτηριακά κύτταρα απομονώνονται με φυγοκέντρηση στις 4.500 - 6.000g, στους 4°C για 10 λεπτά. ακολουθεί εξαγωγή του πλασμιδιακού DNA με χρήση του Plasmid DNA purification kit – Macherey – Nagel Xtra Midi, ακολουθώντας το πρωτόκολλο της κατασκευαστικής εταιρίας ([https://www.mnnet.com/Portals/8/attachments/Redakteure\\_Bio/Protocols/Plasmid%20DNA%20Purification/UM\\_pDNA\\_NuBoXtra.pdf](https://www.mnnet.com/Portals/8/attachments/Redakteure_Bio/Protocols/Plasmid%20DNA%20Purification/UM_pDNA_NuBoXtra.pdf)).

### 3.2.3. Κλωνοποίηση ενθέματος στον pSiCORmod-PGK-puro

Για κάθε καταστολή που πραγματοποιείται, προηγείται κλωνοποίηση μιας αλληλουχίας shRNA στον φορέα pSiCORmod. Καθένα από τα shRNA που χρησιμοποιούνται έχει ως στόχο μια περιοχή στην αλληλουχία του γονιδίου στόχου. Η διαδικασία της κλωνοποίησης περιλαμβάνει τρία στάδια: μια αντίδραση πέψης με ένζυμα περιορισμού, ώστε ο φορέας κλωνοποίησης να αποχωριστεί ένα τμήμα της αλληλουχίας του και να αποκτήσει προεξέχοντα άκρα, την αναδιάταξη των ολιγονουκλεοτιδίων των shRNA που θα σχηματίσουν το δίκλωνο ένθεμα και μια αντίδραση λιγάσης για την ενσωμάτωση του ενθέματος στον κομμένο φορέα.

#### Πέψη με ένζυμα περιορισμού (Digestion reaction)

Για την πέψη του άκοπου pSiCORmod χρησιμοποιούνται ταυτόχρονα κατάλληλα ένζυμα περιορισμού, για τη δημιουργία των προεξέχοντων άκρων. Η αντίδραση διαρκεί συνολικά 5 ώρες, και έχει την εξής σύσταση:

Αντιδραστήρια	Ποσότητες (Ανά δείγμα)
10x Buffer K	5μl
Res. Enzymes – BamHi/XbaI (Minotech)	4μl (2μl +2μl)
Plasmid DNA	1μl
H <sub>2</sub> O	40μl
<b>Τελικός όγκος</b>	<b>50μl</b>

### Αναδιάταξη των ολιγονουκλεοτιδίων shRNA (shRNA oligos annealing)

Τα ολιγονουκλεοτίδια σχηματίζουν μια δίκλωνη μορφή με μία νοηματική (sense) και μια αντινοηματική (antisense) αλυσίδα. Κάθε αλυσίδα περιέχει μία θέση ενζύμου περιορισμού, συμπληρωματική με ένα άκρο του κομμένου φορέα, ώστε να μπορέσει να συνδεθεί με αυτόν μέσω συμπληρωματικότητας. Τα ολιγονουκλεοτίδια περιλαμβάνουν επίσης στην αλληλουχία τους και μια περιοχή σχηματισμού βρόχου, η οποία συμβάλλει στον υβριδισμό νοηματικής και αντινοηματικής αλληλουχίας, ώστε να προκύψει το siRNA που θα προκαλέσει την καταστολή του γονιδίου στόχου. Τα ολιγονουκλεοτίδια βρίσκονται σε λυοφιλιωμένη μορφή και επαναδιαλύονται σε ddH<sub>2</sub>O για δημιουργία stock τελικής συγκέντρωσης 100 μM. Από το συγκεκριμένο stock στη συνέχεια θα παρασκευαστεί το working solution, το οποίο περιέχει 20 μl από το μείγμα των ολιγονουκλεοτιδίων (10μl sense, 10μl antisense) και 180μl ddH<sub>2</sub>O ώστε η τελική συγκέντρωση να ισούται με 10 μM. Στη συνέχεια, ακολουθεί ανάμειξη 2 μl sense (100 μM) και 2μl antisense (100 μM) με 96 μl annealing buffer (50mM NaCl για την ευκολότερη αναδιάταξη των ολιγονουκλεοτιδίων, 10mM Tris, pH=8, για τη ρύθμιση της σύστασης, 1mM EDTA, pH=8, για την παρεμπόδιση της λειτουργίας των ενζύμων). Το μείγμα τοποθετείται σε δοχείο ζέσεως με νερό σε θερμοκρασία 95°C για 5 λεπτά, ώστε να επιτευχθεί αποδιάταξη των δευτεροταγών δομών που σχηματίζουν τα ολιγονουκλεοτίδια. Έπειτα, τοποθετείται σε θερμοκρασία δωματίου για περίπου 1:30 ώρα, προκειμένου να πραγματοποιηθεί ο σταδιακός υβριδισμός και ο σχηματισμός των δίκλωνων ενθεμάτων. Όταν η θερμοκρασία φτάσει στους 30°C το μείγμα (annealed oligos) θα είναι πλέον κατάλληλο για χρήση ως συστατικό στην αντίδραση λιγάσης.

### Αντίδραση λιγάσης (Ligation reaction)

Η αντίδραση λιγάσης πραγματοποιείται με σκοπό την επιτυχή ενσωμάτωση του ενθέματος στον κομμένο φορέα. Το ένθεμα και ο φορέας υβριδίζουν εξαιτίας της συμπληρωματικότητας των βάσεων στις θέσης πέψης από τα ένζυμα περιορισμού. Εάν ο υβριδισμός είναι επιτυχημένος, η λιγάση μπορεί να προσεγγίσει τις θέσεις πέψης και να σχηματίσει τους φωσφορδιστερικούς δεσμούς που λείπουν μεταξύ ενθέματος και φορέα. Με αυτόν τον τρόπο προκύπτει το τελικό προϊόν της αντίδρασης το οποίο θα αξιοποιηθεί για τον μετασχηματισμό, η οποία διαρκεί περίπου 1 με 3 ώρες. Η σύσταση της αντίδρασης είναι η εξής:

Αντιδραστήρια	Ποσότητες (Ανά δείγμα)
ds Oligos (annealed)	10 μl
Digested vector (pSiCORmod) (50 ng/μl)	1,5μl
Buffer (T4 Ligase) 10X	2 μl
T4 DNA Ligase (NEB)	2 μl
H <sub>2</sub> O	4,5 μl
<b>Τελικός όγκος</b>	<b>20μl</b>

Με την ολοκλήρωση της αντίδρασης λιγάσης, το δείγμα οδηγείται για εξαγωγή του πλασμιδιακού DNA και μέτρηση της συγκέντρωσής του μέσω Nanodrop (Quawell). Η εξαγωγή του πλασμιδιακού DNA πραγματοποιήθηκε με τη βοήθεια του Macherey - Nagel Plasmid DNA purification kit, ακολουθώντας το πρωτόκολλο της κατασκευαστικής εταιρίας ([https://www.mnnet.com/Portals/8/attachments/Redakteure\\_Bio/Protocols/Plasmid%20DNA%20Purification/UM\\_pDNA\\_NuBoXtra.pdf](https://www.mnnet.com/Portals/8/attachments/Redakteure_Bio/Protocols/Plasmid%20DNA%20Purification/UM_pDNA_NuBoXtra.pdf)). Ενδεικτικά, στις απομονώσεις που διεξήχθησαν στα πλαίσια των συγκεκριμένων πειραμάτων, η συγκέντρωση του

πλασμιδιακού DNA ανήλθε στα 3159,1 ng/μl. Ως μάρτυρας αρνητικού ελέγχου χρησιμοποιήθηκε δείγμα στο οποίο δεν έχει προστεθεί ένθεμα (self ligation product). Αντί αυτού στο συγκεκριμένο δείγμα προστέθηκε νερό σε ίδια ποσότητα.

#### 3.2.4. Διαγνωστικές πέψεις πλασμιδιακών φορέων (Diagnostic plasmid digestions)

Οι διαγνωστικές πέψεις πραγματοποιούνται με σκοπό την επαλήθευση της ενσωμάτωσης του επιθυμητού ενθέματος στον φορέα κλωνοποίησης. Όσον αφορά τον πλασμιδιακό φορέα pSiCORmod, η πέψη διεξάγεται μέσω της χρήσης κατάλληλων ενζύμων περιορισμού, με τον σχεδιασμό του πειράματος να παρατίθεται παρακάτω.

Αντιδραστήρια	Ποσότητες (Ανά δείγμα)
10x Buffer K	5μl
Res. Enzymes – BamHi/XbaI (Minotech)	4μl (2+2)
Plasmid DNA	1μl
H <sub>2</sub> O	40μl
<b>Τελικός όγκος</b>	<b>50μl</b>

Για την περαιτέρω επαλήθευση του αποτελέσματος, η αντίδραση διαγνωστικής πέψης αξιοποιεί δύο επιπλέον υποστρώματα πλασμιδιακού DNA, τα οποία θα χρησιμοποιηθούν κατά την ηλεκτροφόρηση: Ένα αρνητικό και ένα θετικό δείγμα ελέγχου. Ένα πλασμίδιο στο οποίο δεν έχει πραγματοποιηθεί πέψη εξ αρχής με ένζυμα περιορισμού ή ένα πλασμίδιο το οποίο έχει επεξεργαστεί με ένζυμα περιορισμού, ωστόσο το τμήμα πωσ εκτομή δεν έχει αποσπασθεί επιτυχώς και η αρχική αλληλουχία του έχει διατηρηθεί, μπορεί να αποτελέσει ένα αρνητικό δείγμα ελέγχου (Negative Control). Ένα πλασμίδιο στο οποίο έχει επιβεβαιωθεί η ένθεση του επιθυμητού ενθέματος με πολλούς τρόπους (διαγνωστική πέψη, διαγνωστική PCR και αλληλούχηση) μπορεί να αποτελέσει ένα θετικό δείγμα ελέγχου (Positive Control).

#### 3.2.5. Αλληλούχηση

Η αλληλούχηση παρέχει την μέγιστη δυνατή ακρίβεια και την αξιοπιστία για τον έλεγχο της ένθεσης της επιθυμητής αλληλουχίας σε έναν φορέα κλωνοποίησης. Για την επίτευξη της παραπάνω διαδικασίας τα δείγματα παραπέμφθηκαν για αλληλούχηση στο κέντρο της CeMIA (<https://cemia.eu>). Τα δείγματα προετοιμάστηκαν βάσει των προδιαγραφών της ηλεκτρονικής πλατφόρμας του κέντρου εφαρμογών.

#### 3.2.6. Υποκυτταρική Κλασμάτωση

Το πείραμα υποκυτταρικής κλασμάτωσης, δηλαδή της ελεγχόμενης λύσης των κυττάρων και της διαδοχικής απομόνωσης υποκυτταρικών κλασμάτων μέσω διαφορικών φυγοκεντρήσεων, περιλαμβάνει την συλλογή τριών διαφορετικών κλασμάτων. Τα κλάσματα αυτά αντιστοιχούν στο κυτταρόπλασμα (supernant 2 - S2), στο πυρηνόπλασμα (supernant 3 - S3) και στην αδιάλυτη χρωματίνη (pellet 3 - P3). Η διαδικασία ξεκινάει με θρυψινοποίηση της καλλιέργειας των κυττάρων, ακολουθούμενη από προσθήκη θρεπτικού υλικού για αναστολή της δράσης της θρυψίνης. Το κυτταρικό ελαιώρημα που προκύπτει συλλέγεται σε tubes του 1,5 ml. Η συλλογή πραγματοποιείται σε 2 tubes σε συνολικό όγκο 3 ml, με το κάθε tube να αντιστοιχεί σε ένα μελλοντικό κλάσμα. Το αρχικό tube για τη συλλογή των κυττάρων θα αποτελέσει χώρο αποθήκευσης της αδιάλυτης χρωματίνης (P3). Με τη χρήση εργαλείου

Dounce πραγματοποιείται η αποδοτικότερη δυνατή θραύση των κυτταρικών μεμβρανών, ώστε να προκύψει ένα ομογενές διάλυμα κυττάρων. Ακολούθως, πραγματοποιείται φυγοκέντρηση στις 3000rpm, στους 4°C για 5 λεπτά. Με το πέρας της φυγοκέντρησης απορρίπτεται το υπερκείμενο και το ίζημα (P1), επαναδιαλύεται σε Buffer λύσης A, σε αναλογία 1/10. Ακολουθεί δεύτερη φυγοκέντρηση σε αυξημένη ταχύτητα. Το υπερκείμενο που προκύπτει αντιστοιχεί στο κυτταρόπλασμα και συλλέγεται σε νέο tube με όνομα S2. Στα συγκεκριμένα tubes προστίθεται τριζόλη σε ποσότητα ίση με αυτή του απορριφθέντος υπερκειμένου πριν αυτά αποθηκευτούν στους -30°C. Στο ίζημα που προκύπτει μετά την δεύτερη φυγοκέντρηση επαναδιαλύεται με την προσθήκη buffer A και buffer NLB800 σε αναλογία 1:1 μεταξύ τους. Τα tubes μεταφέρονται σε ειδικό ρότορα στους 4°C για 30 λεπτά. Ακολουθεί φυγοκέντρηση σε 14.000 rpm στους 4°C, για 15 λεπτά. Το υπερκείμενο που θα προκύψει αντιστοιχεί στο πυρηνόπλασμα και συλλέγεται σε νέα tubes με όνομα S3. Το ίζημα της αδιάλυτης χρωματίνης παραμένει στα tubes με το όνομα P3. Η προσθήκη τριζόλης επαναλαμβάνεται για τα κλάσματα S3 και P3. Στη συνέχεια, τα δείγματα αποθηκεύονται στους -30°C. Τα buffer που χρησιμοποιήθηκαν είναι υποτονικά και με διαφορά στην μεταξύ τους σύσταση, με σκοπό την επιτυχημένη διάσπαση των κυτταρικών και των πυρηνικών μεμβρανών. Τα παραπάνω δείγματα αξιοποιήθηκαν για απομόνωση RNA, κατεργασία με DNάση, σύνθεση cDNA και ποσοτική PCR (qPCR), διαδικασίες οι οποίες περιγράφονται παρακάτω.

### 3.2.7. Απομόνωση RNA (RNA Extraction)

Η απομόνωση του RNA, δηλαδή η εξαγωγή του RNA από κύτταρα και η διατήρησή του σε καθαρή μορφή ξεκινά με την προσθήκη 1ml τριζόλης σε δείγμα της κυτταρικής σειράς προς μελέτη. Καθένα από τα δείγματα διατηρείται για ένα λεπτό σε θερμοκρασία δωματίου πριν μεταφερθεί στον πάγο. Ακολουθεί προσθήκη 200μl χλωροφορμίου και ανάδευση μέσω vortex για 30 δευτερόλεπτα. Στη συνέχεια, πραγματοποιείται φυγοκέντρηση στις 14.000 rpm, στους 4°C για 5 λεπτά. Μετά το τέλος της φυγοκέντρησης έπεται η μεταφορά του υπερκειμένου σε ξεχωριστό tube συνοδευόμενη από προσθήκη ισοπροπανόλης σε αναλογία 1:1 και 5 μl γλυκογόνου. Το δείγμα τοποθετείται για 10 λεπτά στον πάγο και έπειτα φυγοκεντρείται στις 14.000 rpm, στους 4°C για 15 λεπτά. Με το πέρας της φυγοκέντρησης, το υπερκείμενο απορρίπτεται, και το δείγμα τοποθετείται με ανοιχτό το καπάκι του tube στον πάγο για περίπου 15 λεπτά, προκειμένου να στεγνώσει το ίζημα. Με το πέρας των 15 λεπτών και ανάλογα με το μέγεθος του, το ίζημα επαναδιαλύεται σε όγκο 30 ή 50 μl (ddH<sub>2</sub>O). Ακολουθεί θέρμανση στους 65°C για 5 λεπτά, ανάδευση μέσω vortex και αποθήκευση στους -80°C.

### 3.2.7. Κατεργασία με DNάση (DNase treatment)

Η κατεργασία με τη χρήση DNάσης έπεται του πρωτοκόλλου απομόνωσης του RNA και περιλαμβάνει την επώαση του δείγματος με μια δεοξυριβουνοκλεάση, η οποία καταστρέφει τα μόρια του DNA χωρίς να καταστρέφει το RNA. Με την ολοκλήρωση της απομόνωσης του RNA το δείγμα οδηγείται για μέτρηση της συγκέντρωσής του μέσω Quawell. Στη συνέχεια, προετοιμάζεται MIX με την ακόλουθη σύσταση:

MIX	Ποσότητες (Ανά δείγμα)
Buffer 10x	5 μl
DNase (Thermofisher)	2 μl
RNaseout	0,5 μl
ddH <sub>2</sub> O	2,5 μl
<b>Τελικός όγκος</b>	<b>10 μl</b>

Όσον αφορά την απομόνωση RNA που πραγματοποιήθηκε για τον προσδιορισμό των επιπέδων καταστολής των γονιδίων στόχων, το συγκεκριμένο MIX προστέθηκε σε ιζήμα το οποίο επαναδιαλύθηκε σε όγκο 50 μl. Και τα 50 μl χρησιμοποιήθηκαν για την κατεργασία με την DNάση. Επομένως, ο συνολικός όγκος της αντίδρασης ανήλθε στα 60 μl. Η ίδια επαναδιάλυση του στο πείραμα μελέτης της μεταγραφικής σταθερότητας πραγματοποιήθηκε στα 25 μl.

Ακολουθεί vortex, spindown και τοποθέτηση σε υδατόλουτρο για μία ώρα στους 37°C. Με το πέρας της μίας ώρας, προστίθενται 40 μl οξικού νατρίου (3M, pH=5,6) και 310 μl ddH<sub>2</sub>O. Η διαδικασία συνεχίζεται με την προσθήκη 400 μl φαινόλης/χλωροφορμίου, vortex και φυγοκέντρηση στις 14.000 rpm, στους 4°C για 5 λεπτά. Στο εν λόγω στάδιο, κατά τη μελέτη των επιπέδων καταστολής των γονιδίων στόχων, χορηγήθηκαν 550 μl φαινόλης/χλωροφορμίου. Το υπερκείμενο που προκύπτει μετά τη φυγοκέντρηση συλλέγεται σε νέο tube. Στο δείγμα προστίθεται 1 ml παγωμένης αιθανόλης 100%, πραγματοποιείται ελαφρά ανακίνηση και διατήρηση του στους -80°C για 1 ώρα. Σε αυτό το στάδιο προστίθενται και 10 μl γλυκογόνου. Με τη λήξη της μίας ώρας το δείγμα φυγοκεντρείται στις 14.000 rpm, στους 4°C για 30 λεπτά. Το υπερκείμενο απορρίπτεται και αντικαθίσταται από 500 μl παγωμένης αιθανόλης 70% και φυγοκεντρείται εκ νέου στις 14.000 rpm, στους 4°C για 5 λεπτά. Το νέο υπερκείμενο αφαιρείται προσεκτικά για την αποτροπή της διατάραξης του ιζήματος και τοποθετείται στον πάγκο για 10 λεπτά προκειμένου να στεγνώσει. Ακολουθεί επαναδιάλυση του ιζήματος σε 30μl ddH<sub>2</sub>O, θέρμανση σε υδατόλουτρο στους 65°C για 5 λεπτά, vortex για 30 δευτερόλεπτα και spindown. Στην προκειμένη περίπτωση, στα δείγματα μελέτης της καταστολής, η επαναδιάλυση του ιζήματος πραγματοποιήθηκε στα 15 μl ddH<sub>2</sub>O. Μετά από αυτή τη διαδικασία, τα δείγματα μπορούν πλέον να αποθηκευτούν στους -80°C. Σε αυτό το στάδιο υπάρχει δυνατότητα της μέτρησης της συγκέντρωσης του RNA του δείγματος μέσω Quawell.

### 3.2.8. Σύνθεση cDNA (cDNA synthesis - Reverse Transcription)

Η σύνθεση του cDNA ή αλλιώς η αντίδραση αντίστροφης μεταγραφής περιλαμβάνει την σύνθεση μορίων cDNA από τα δείγματα απομόνωσης RNA, μέσω ενός ενζύμου που ονομάζεται αντίστροφη μεταγραφάση (MMLV). Τα δείγματα προετοιμάζονται κατάλληλα ώστε να περιέχονται 1000ng διαλυμένα σε 10μl ddH<sub>2</sub>O. Για την διαδικασία της κατεργασίας χρησιμοποιήθηκαν δύο MIX, η σύσταση των οποίων παρατίθεται στους παρακάτω πίνακες.

MIX 1	Ποσότητες (Ανά δείγμα)
OligodT	1 μl
dNTPs	1 μl
<b>Τελικός όγκος</b>	<b>2 μl</b>

MIX 2	Ποσότητες (Ανά δείγμα)
5X Buffer	4 μl
100nm DTT	2 μl
Rnase out	1 μl
MMLV (Reverse Transcriptase) (Invitrogen)	1 μl
<b>Τελικός όγκος</b>	<b>8 μl</b>

Σε δείγμα RNA όγκου 10 μl προστίθενται 2 μl από το MIX 1, ακολουθεί επώαση σε υδατόλουτρο στους 65°C για 5 λεπτά και στη συνέχεια 8 μl από το MIX 2 με τον τελικό όγκο να ανέρχεται στα 20 μl. Το τελικό μείγμα τοποθετείται σε υδατόλουτρο στους 37°C για 2 ώρες. Με το πέρας των δύο ωρών, πραγματοποιείται αραιώση του δείγματος σε 380μl ddH<sub>2</sub>O. Σε αυτό το στάδιο, τα δείγματα μπορούν να αποθηκευτούν σε θερμοκρασία -20°C.

### 3.2.9. Ποσοτική PCR πραγματικού χρόνου (Real time quantitative PCR - qPCR)

Η qPCR αποτελεί μια μέθοδο προσδιορισμού των επιπέδων έκφρασης γονιδίων προς μελέτη στο αρχικό δείγμα. Η μέθοδος αξιοποιεί την ικανότητα της χρωστικής SYBR green να προσδένεται ισχυρά στο δίκλωνο DNA εκπέμποντας ισχυρότερο φθορισμό σε σχέση με την πρόσδεσή του στο μονόκλωνο DNA. Αυτή η ιδιότητα της χρωστικής αξιοποιείται, ώστε να επιτυγχάνεται η ποσοτικοποίηση των γονιδίων στόχων που ενισχύονται με τη βοήθεια ειδικών εκκινητών οι οποίοι υβριδίζουν στα επιθυμητά μόρια cDNA. Η ποσοτικοποίηση πραγματοποιείται στην εκθετική φάση της καμπύλης ενίσχυσης, κατά την οποία τα προϊόντα διπλασιάζονται σε αριθμό και ξεπερνούν ένα κατώφλι φθορισμού, το οποίο αντιστοιχεί σε έναν συγκεκριμένο κύκλο διπλασιασμού (C<sub>t</sub>). Παράλληλα με την ενίσχυση του γονιδίου στόχου στα δείγματα, πραγματοποιείται και ενίσχυση μιας γονιδιωματικής περιοχής σε δείγματα ανθρώπινου γονιδιωματικού DNA. Τα δείγματα αυτά έχουν γνωστή συγκέντρωση και έχουν παρασκευαστεί κατάλληλα ώστε να έχουν υποδεκαπλάσιες αραιώσεις. Τα συγκεκριμένα δείγματα χρησιμοποιούνται ως standards με σκοπό την κατασκευή πρότυπης καμπύλης αναφοράς, η οποία αξιοποιείται για τον προσδιορισμό του C<sub>t</sub> και την ποσοτικοποίηση. Επιπλέον, για την κανονικοποίηση της έκφρασης των γονιδίων στόχων χρησιμοποιούνται δύο ιδιοσυστάτα γονίδια, το GAPDH και το RPLP1. Η προετοιμασία και η οργάνωση των δειγμάτων, καθώς και ο προγραμματισμός των πρωτοκόλλων για την εκτέλεση της qPCR πραγματοποιήθηκαν με το πρόγραμμα CFX manager της εταιρίας BioRad. Τα δείγματα της qPCR προετοιμάζονται σε ειδικά tubes σε strips με την σύσταση που παρατίθεται παρακάτω:

Αντιδραστήρια	Ποσότητες (Ανά δείγμα)
SYBR green Buffer Mix 2x	10 μl
Forward/Reverse primers	2 μl
Template (cDNA)	4 μl
H <sub>2</sub> O	4 μl
<b>Τελικός όγκος</b>	<b>20 μl</b>

Για την επαλήθευση της αξιοπιστίας της αντίδρασης, χρησιμοποιούνται επίσης δείγματα αρνητικού ελέγχου. Πιο συγκεκριμένα, χρησιμοποιείται ένα δείγμα απ' όπου απουσιάζει το δείγμα cDNA (No Template Control - NTC) καθώς και ένα δείγμα στο οποίο δεν έχει πραγματοποιηθεί αντίδραση αντίστροφης μεταγραφής (No Reverse Transcriptase - NRT).

### 3.2.10. Πείραμα ελέγχου μεταγραφικής σταθερότητας (Transcriptional stability assay)

Το πείραμα μεταγραφικής σταθερότητας εξυπηρετεί την παρατήρηση ενός φαινομένου σε διαφορετικές χρονικές στιγμές (timepoints) στο πλαίσιο μιας χρονικής περιόδου. Σκοπός του εν λόγω πειράματος ήταν ο προσδιορισμός του χρόνου υποδιπλασιασμού του αριθμού των μεταγράφων στόχων σε καρκινικές κυτταρικές σειρές στομάχου, σε συνάρτηση με τον χρόνο έκθεσης τους στην ακτινομυκίνη. Πιο συγκεκριμένα, στο πείραμα που διεξήχθη στα πλαίσια της παρούσας διπλωματικής εργασίας, οι χρονικές στιγμές ήταν οι 0h, 0,5h, 1h, 2h, 4h, 6h, 8h και 24h ενώ χρησιμοποιήθηκε η καρκινική κυτταρική σειρά στομάχου NUGC4. Η ακτινομυκίνη δρα ως καταστολέας της μεταγραφής για

τα κύτταρα, παρουσιάζοντας χρονοεξαρτώμενη δράση και δίνοντας την δυνατότητα προσδιορισμού του ρυθμού αποικοδόμησης των μεταγράφων. Στο συγκεκριμένο πείραμα η συγκέντρωση ακτινομυκίνης που χορηγήθηκε ήταν 2  $\mu\text{g}/\text{ml}$ . Το δείγμα στην χρονική στιγμή 0h χρησιμοποιείται ως δείγμα αναφοράς για την έναρξη του πειράματος. Αποτελεί δείγμα κυττάρων απουσία ακτινομυκίνης στο οποίο έχει προστεθεί τριζόλη με σκοπό τη μονιμοποίηση των κυττάρων και την προστασία του RNA. Μια μέρα πριν την έναρξη του πειράματος, προετοιμάζεται ένα τρυβλίο για κάθε timepoint. Τα κύτταρα αποδεσμεύονται από τα τοιχώματα της φλάσκας μέσω της προσθήκης θρυψίνης. Στη συνέχεια χορηγείται ποσότητα θρεπτικού για την απενεργοποίηση της δράσης της. Ακολουθεί ανάμειξη 10  $\mu\text{l}$  από κύτταρα με 90  $\mu\text{l}$  χρωστικής trypan blue. Από το μείγμα αυτό, 10  $\mu\text{l}$  τοποθετούνται σε ειδική πλάκα Neubauer, ώστε να προσδιοριστεί ο αριθμός των κυττάρων του δείγματος. Έπειτα, τα κύτταρα προστίθενται στα τρυβλία παρουσία θρεπτικού, σε τελικό όγκο 3ml και τοποθετούνται σε επωαστήρα στους 37°C, σε 5% CO<sub>2</sub>. Μεταξύ των διαστημάτων πραγματοποιείται απόρριψη του θρεπτικού υλικού από τα τρυβλία και εκ νέου χορήγηση φρέσκου θρεπτικού. Μετά το πέρας των χρονικών διαστημάτων προστίθεται τριζόλη σε κάθε δείγμα και ακολουθεί αποθήκευση στους -80°C. Τη διαδικασία ακολουθούν απομόνωση RNA κατεργασία με DNάση, σύνθεση cDNA και qPCR. Ως δείκτης αξιοπιστίας του πειράματος χρησιμοποιήθηκε το γονίδιο MYC, το οποίο παρουσιάζει χρόνο υποδιπλασιασμού 30 λεπτά και παρουσιάζει ιδιαίτερη ευαισθησία όσον αφορά την σταθερότητα των μεταγράφων του.

### 3.2.11. Δοκιμασία ελέγχου σχηματισμού αποικιών (Colony Formation Assay - CFA)

Σκοπός της δοκιμασίας είναι ο προσδιορισμός του δυναμικού πολλαπλασιασμού των κυττάρων και κατά συνέπεια του φαινοτύπου τους, σε μακροσκοπικό επίπεδο. Για το σκοπό αυτό, NUGC4 κύτταρα ( $3 \times 10^5$ ) τα οποία έγινε υπερέκφραση shRNA είτε έναντι των CENTRIC1/2 είτε με αλληλουχίες control (scrambled) αναπτύχθηκαν σε six well plates παρουσία RPMI/10% FBS θρεπτικού (Gibco) στους 37°C / 5% CO<sub>2</sub>. Έπειτα από 6 ημέρες αφαιρέθηκε το θρεπτικό μέσο και τα κύτταρα ξεπλύθηκαν 1 φορά με 1x PBS. Ακολούθως προστέθηκε για 5 λεπτά διάλυμα κουμάσης 0.25% σε 45/45/10 Μεθανόλη, νερό, οξικό οξύ και έπειτα ακολούθησαν 2 εκπλύσεις, πρώτα με 50/50 Μεθανόλη/νερό και έπειτα άλλα 2 μόνο με νερό. Αφού αφέθηκαν να στεγνώσουν για 24 ώρες, τα πιάτα φωτογραφήθηκαν σε στερεοσκόπιο.

## 4. ΑΠΟΤΕΛΕΣΜΑΤΑ

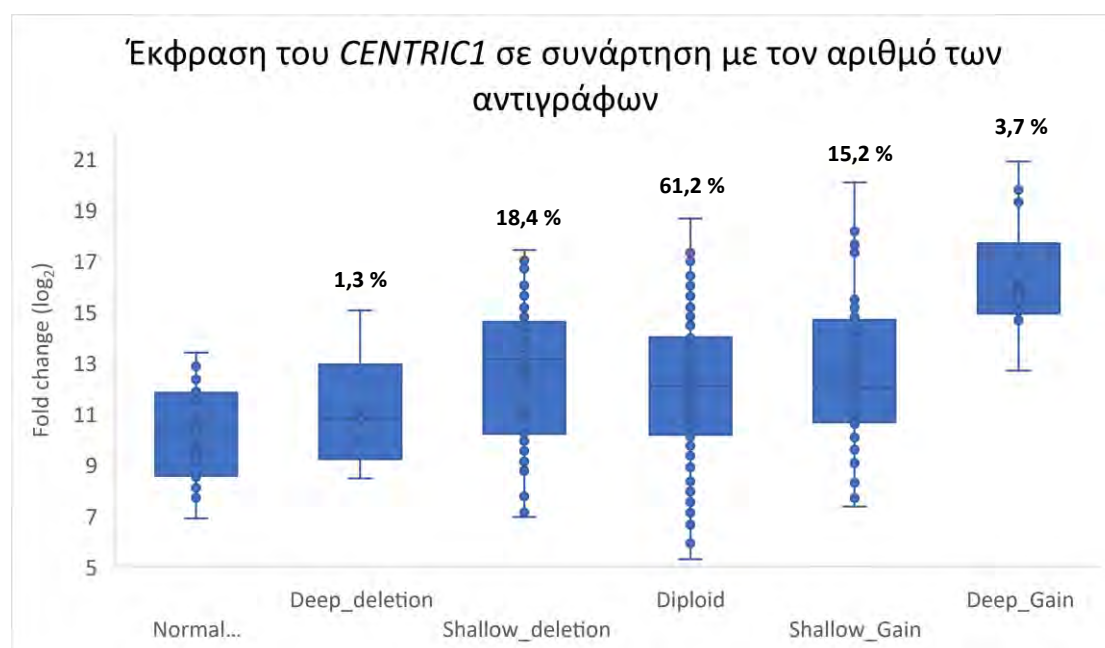


## 4.1 Βιοπληροφορική Ανάλυση

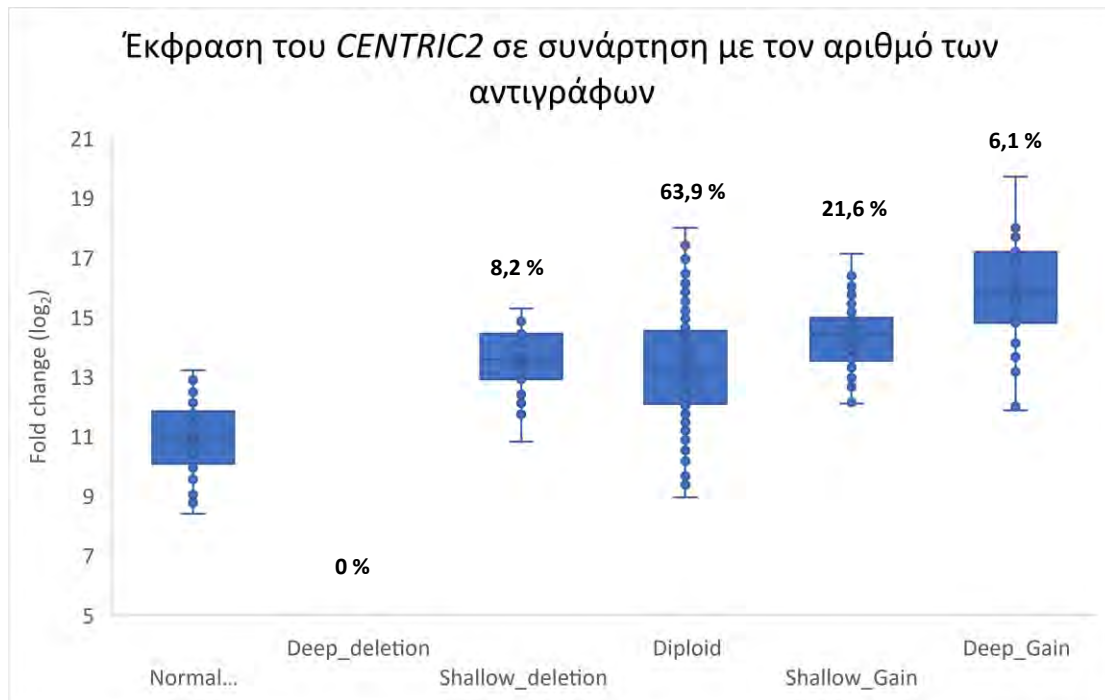
### 4.1.1 Εντοπισμός γονιδίων με παραλλαγές αριθμού αντιγράφων

Προκειμένου να μελετηθεί η αλληλεπίδραση μεταξύ lncRNAs και CNVs, το πρώτο βήμα ήταν ο προσδιορισμός γονιδίων στόχων, τα οποία πληρούν την επιθυμητή θετική συσχέτιση του αυξημένου αριθμού αντιγράφων με αυξημένη γονιδιακή έκφραση. Γι' αυτό, πραγματοποιήθηκε μετανάλυση δεδομένων RNA-seq σε δείγματα βιοψιών γαστρικού καρκίνου. Τα δείγματα αυτά αφορούν την έκφραση γονιδίων σε περιπτώσεις γονιδιωμάτων που φέρουν αριθμό αντιγράφων σε φυσιολογική διπλοειδή μορφή (Diploid), μειωμένο αριθμό αντιγράφων ενός (Shallow deletion) ή δύο αλληλομόρφων (Deep deletion) καθώς και αυξημένο αριθμό αντιγράφων (1-2 - Shallow gain) ή πολύ αυξημένο αριθμό αντιγράφων (>2 - Deep gain). Επιπλέον, παρατίθενται και δεδομένα έκφρασης από βιοψίες παρακαρκινικών ιστών στομάχου (Normal Paracancerous) στους οποίους δεν έχουν παρατηρηθεί ιστολογικές αλλοιώσεις. Τα διαγράμματα της έκφρασης παρουσιάζονται σε μορφή box plot.

Με βάση τα παραπάνω, για την μελέτη στα πλαίσια της παρούσας πτυχιακής επιλέχθηκαν δύο γονίδια, τα *CENTRIC1* και *CENTRIC2*. Πιο συγκεκριμένα, τα δύο αυτά γονίδια παρουσιάζουν αυξημένη έκφραση όσο αυξάνεται ο αριθμός των αντιγράφων τους στο γονιδίωμα. Τα αποτελέσματα των αναλύσεων συσχέτισης μεταξύ έκφρασης και αριθμού αντιγράφων παρατίθενται παρακάτω (**Εικόνα 15A,B**). Στη συνέχεια, παρατίθενται περαιτέρω βιοπληροφορικές αναλύσεις με ενδεικτικό παράδειγμα το γονίδιο *CENTRIC2*. Τα αποτελέσματα ισχύουν αντίστοιχα και για το *CENTRIC1*.



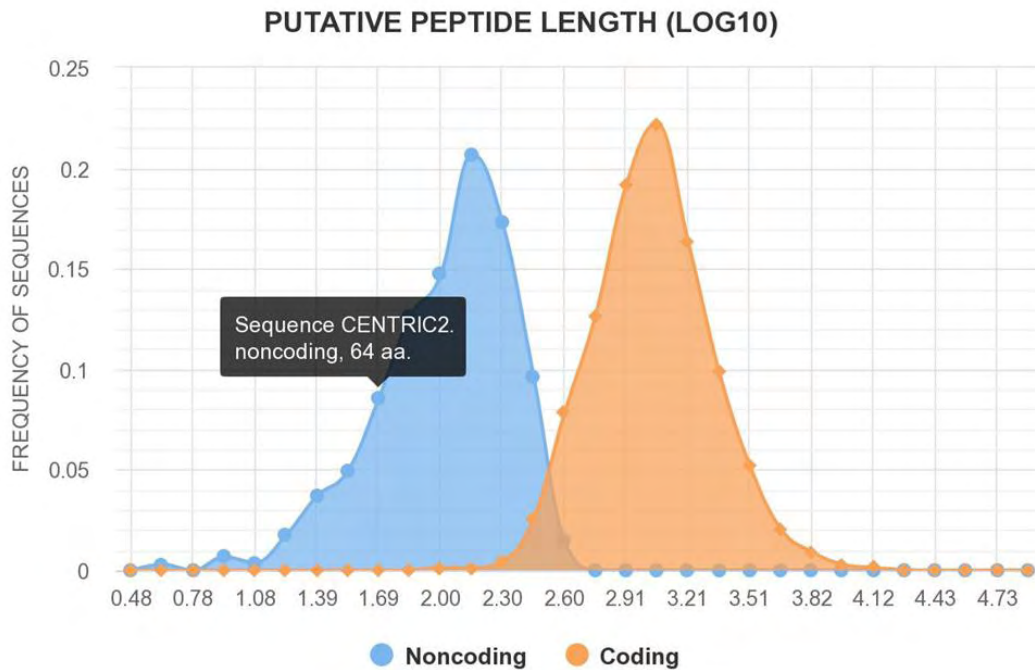
**Εικόνα 15A:** Διάγραμμα συσχέτισης της έκφρασης με τον αριθμό των γονιδιακών αντιγράφων για το γονίδιο *CENTRIC1*. Η έκφραση φαίνεται να αυξάνεται όσο αυξάνεται ο αριθμός των αντιγράφων στο γονιδίωμα. Τα ποσοστά υποδηλώνουν ξεχωριστά τον επί τοις εκατό αριθμό των βιοψιών για κάθε γονιδιακή δόση, ως προς τον σύνολο των καρκινικών βιοψιών που αξιοποιήθηκαν για την ανάλυση.



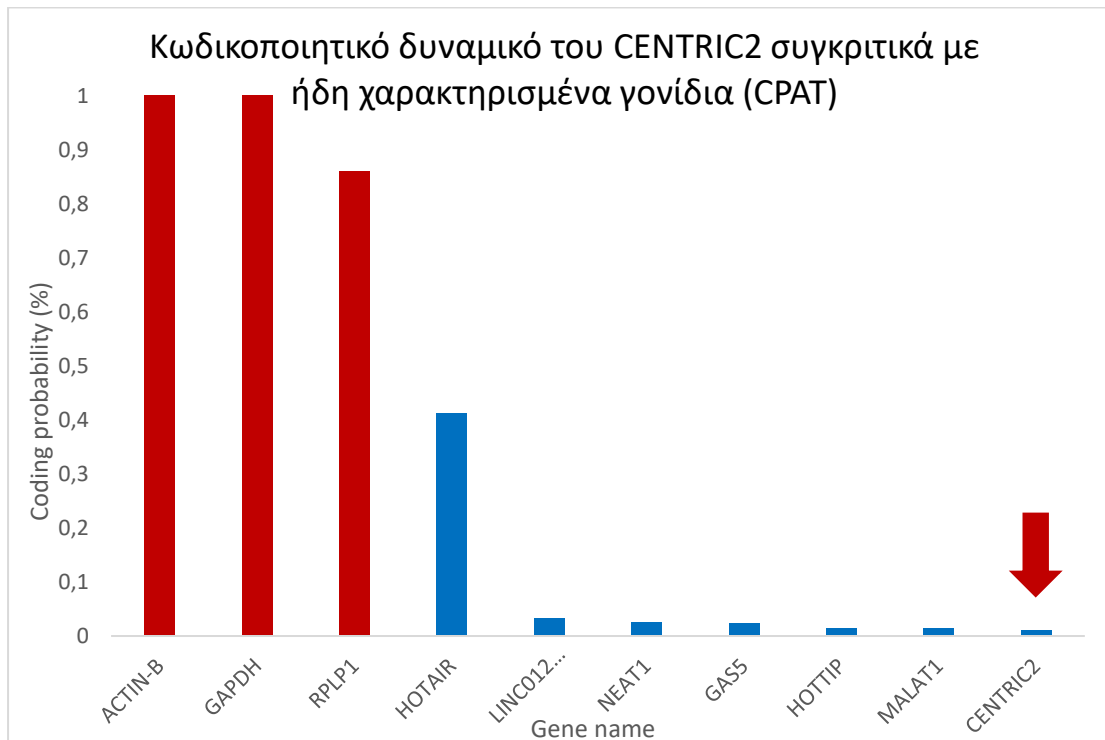
**Εικόνα 15B:** Διάγραμμα συσχέτισης της έκφρασης με τον αριθμό των γονιδιακών αντιγράφων για το γονίδιο *CENTRIC2*. Η έκφραση φαίνεται να αυξάνεται όσο αυξάνεται ο αριθμός των αντιγράφων στο γονιδίωμα. Τα ποσοστά υποδηλώνουν ξεχωριστά τον επί τοις εκατό αριθμό των βιοψιών για κάθε γονιδιακή δόση, ως προς τον σύνολο των καρκινικών βιοψιών που αξιοποιήθηκαν για την ανάλυση.

#### 4.1.1 Επαλήθευση κωδικοποιητικού δυναμικού

Προκειμένου να επιβεβαιωθεί ότι τα *CENTRIC1* και *CENTRIC2* αποτελούν μη κωδικοποιητικά lncRNA μετάγραφα, διεξήχθησαν δοκιμασίες πρόβλεψης κωδικοποιητικού δυναμικού για τα δύο γονίδια μέσω δύο προγραμμάτων, του CPC2 (**Εικόνα 16A**) και του CPAT (**Εικόνα 16B**). Πιο συγκεκριμένα, τα δύο αυτά προγράμματα υπολογίζουν την πιθανότητα ενός γονιδίου να φέρει στην αλληλουχία του έναν ανοιχτό πλαίσιο ανάγνωσης που μπορεί να κωδικοποιήσει για μια πρωτεΐνη ή ένα πολυπεπίδιο, εξάγοντας μια αντιπροσωπευτική τιμή (Fickett score). Ως γονίδια ελέγχου της δοκιμασίας χρησιμοποιήθηκαν καλά χαρακτηρισμένα κωδικοποιητικά (*ACTIN-B*, *GAPDH*, *RPLP1*) αλλά και μη κωδικοποιητικά γονίδια (*HOTAIR*, *NEAT1*, *GAS5*, *HOTTIP*, *MALAT1*, *LINC01234*). Το κωδικοποιητικό δυναμικό των παραπάνω γονιδίων υπολογίστηκε μαζί με το γονίδιο *CENTRIC2* και στη συνέχεια όλα τα γονίδια τοποθετήθηκαν σε διάγραμμα για την καλύτερη δυνατή σύγκριση. Τα αποτελέσματα και των δύο προγραμμάτων επιβεβαίωσαν ότι το γονίδιο *CENTRIC2* αποτελεί ένα μη κωδικοποιητικό lncRNA.



**Εικόνα 16A:** Υπολογισμός κωδικοποιητικού δυναμικού του γονιδίου *CENTRIC2* μέσω του προγράμματος CPC2. Στο διάγραμμα απεικονίζεται η ταξινόμηση του γονιδίου με βάση το Fickett score. Τα γονίδια με κωδικοποιητικό δυναμικό αναμένονται εντός του πορτοκαλί πλαισίου, ενώ τα γονίδια με μη κωδικοποιητικό δυναμικό εντός του γαλάζιου πλαισίου.

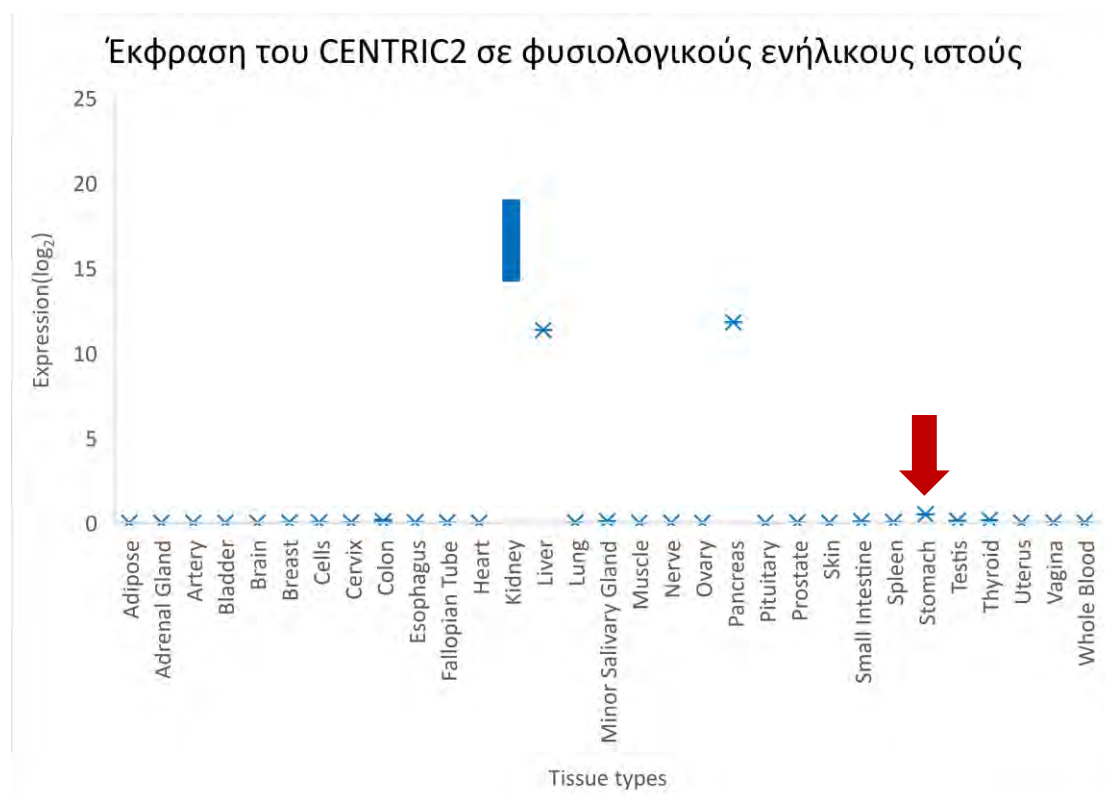


**Εικόνα 16B:** Υπολογισμός κωδικοποιητικού δυναμικού του γονιδίου *CENTRIC2* μέσω του προγράμματος CPAT. Με κόκκινο χρώμα απεικονίζονται τα γονίδια με κωδικοποιητικό δυναμικό, των οποίων οι τιμές ανέρχονται κοντά στο 1. Με μπλε χρώμα απεικονίζονται τα μη κωδικοποιητικά μεταγγραφα με τιμές μικρότερες του 0,5, καθώς και το *CENTRIC2*, επιβεβαιώνοντας την ιδιότητά του ως ένα μη κωδικοποιητικό lncRNA.

## 4.1.2 Ανάλυση του προφίλ έκφρασης του lncRNA CENTRIC2

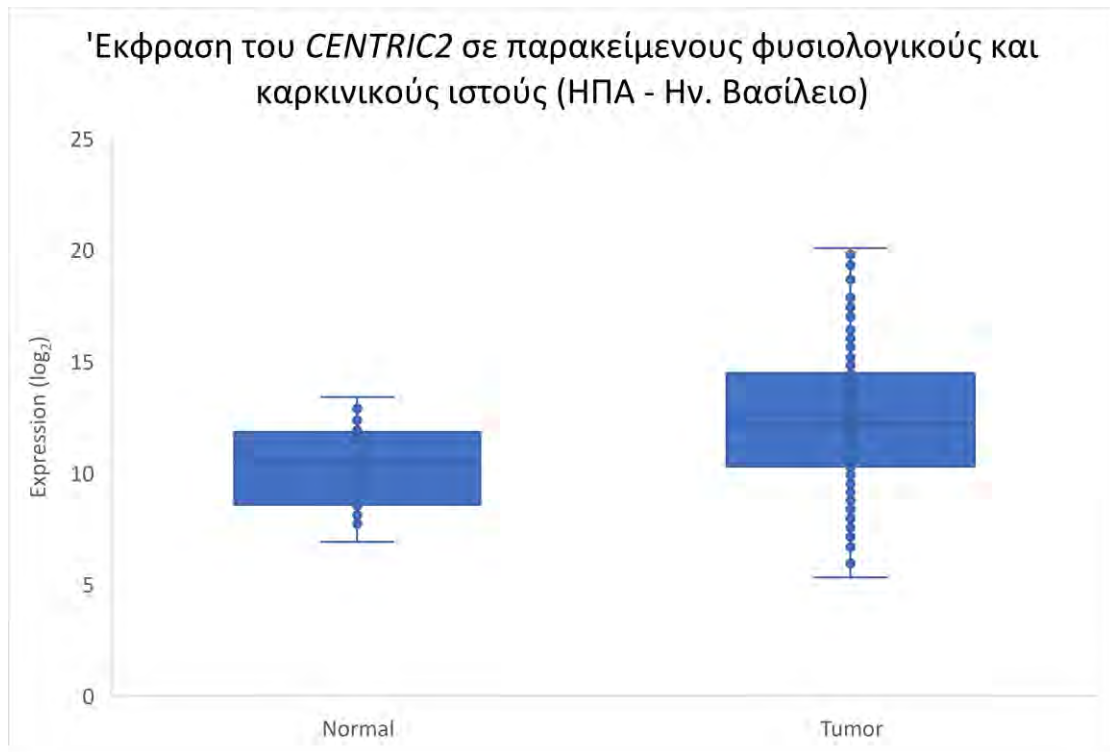
### 4.1.2.1 Έκφραση του CENTRIC2 σε δείγματα βιοψιών

Επόμενο βήμα για τον χαρακτηρισμό του CENTRIC2 ήταν η ανάλυση της έκφρασης του lncRNA, προκειμένου να προσδιοριστεί ένα αρχικό προφίλ έκφρασης του γονιδίου. Αρχικά, πραγματοποιήθηκε μετανάλυση δεδομένων RNAseq σε 55 δείγματα βιοψιών φυσιολογικών ενήλικων ιστών με βάση τα αποτελέσματα της κοινοπραξίας gTEX. Τα αποτελέσματα της έκφρασης παρατίθενται παρακάτω (**Εικόνα 17**).

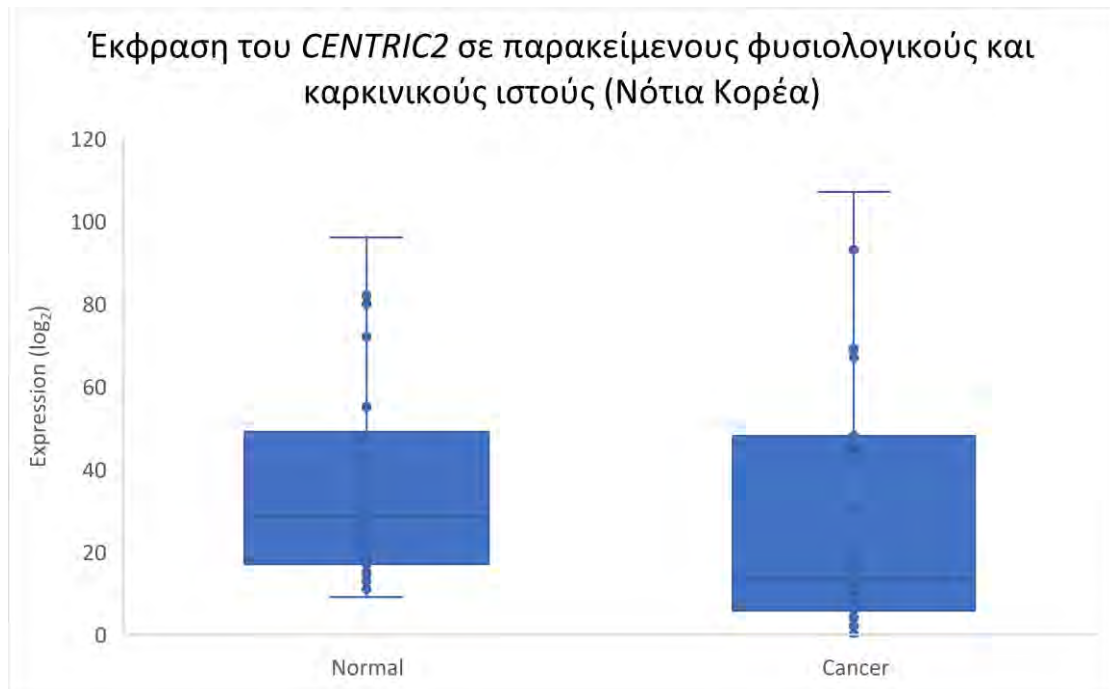


**Εικόνα 17:** Έκφραση του lncRNA CENTRIC2 σε δείγματα βιοψιών φυσιολογικών ενήλικων ιστών. Πηγή: <https://www.biorxiv.org/content/10.1101/793091v1>

Ακολούθησε ανάλυση της έκφρασης του CENTRIC2 σε καρκινικές βιοψίες στομάχου σε αντιπαράθεση με παρακαρκινικές γαστρικές βιοψίες. Γι' αυτό το σκοπό αξιοποιήθηκαν 408 δείγματα καρκινικού και παρακείμενου φυσιολογικού ιστού από άτομα με καταγωγή από το Ηνωμένο Βασίλειο και τις ΗΠΑ (**Εικόνα 18A**), καθώς και 69 δείγματα προερχόμενα από την Νότια Κορέα (**Εικόνα 18B**) (TCGA consortium).

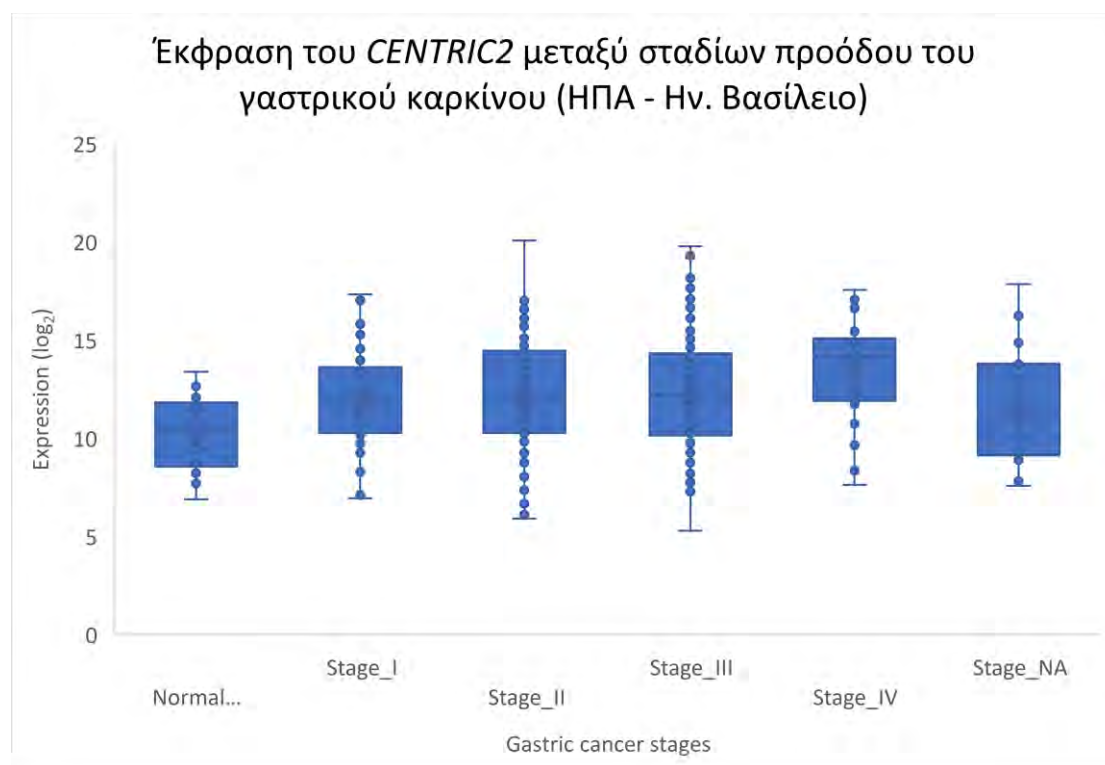


Εικόνα 18Α: Ανάλυση της έκφρασης του *CENTRIC2* σε δείγματα καρκινικού και παρακείμενου φυσιολογικού ιστού ασθενών από τις ΗΠΑ και το Ηνωμένο Βασίλειο.



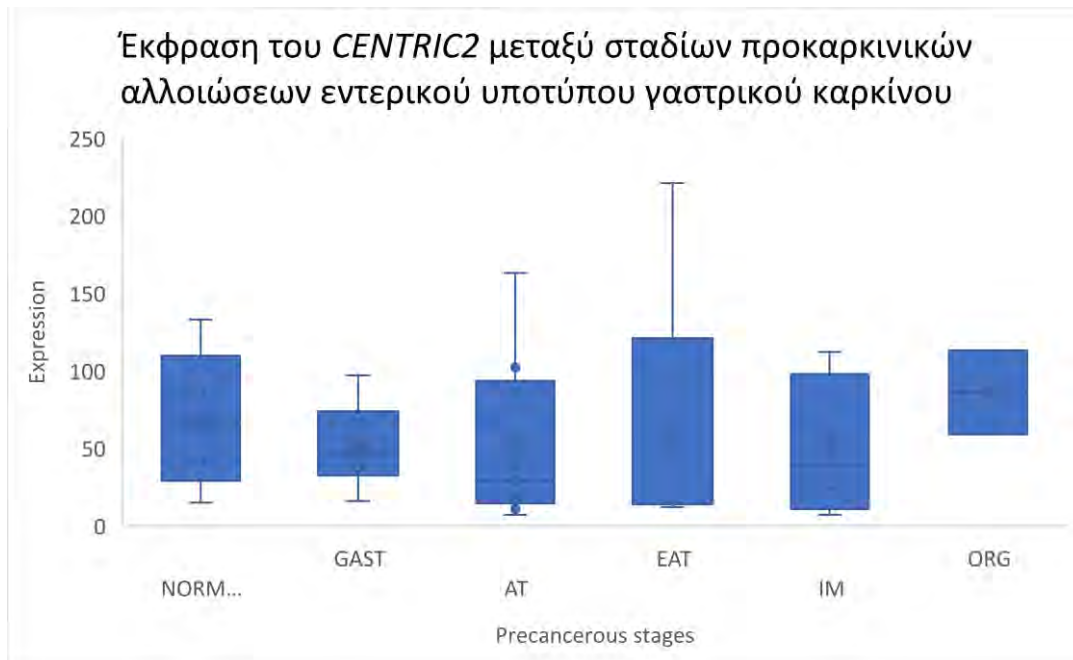
Εικόνα 18Β: Ανάλυση της έκφρασης του *CENTRIC2* σε δείγματα καρκινικού και παρακείμενου φυσιολογικού ιστού ασθενών από την Νότια Κορέα.

Η ίδια ανάλυση πραγματοποιήθηκε και σε επίπεδο καρκινικών σταδίων. Πιο συγκεκριμένα, διεξήχθη μετανάλυση δεδομένων RNAseq σε δείγματα βιοψιών ανά καρκινικό στάδιο γαστρικού καρκίνου. Τα δεδομένα που αξιοποιήθηκαν αντιστοιχούν σε 408 δείγματα βιοψιών καρκινικού και παρακείμενου φυσιολογικού ιστού στομάχου από άτομα με καταγωγή από το Ηνωμένο Βασίλειο και τις ΗΠΑ (**Εικόνα 19**).



**Εικόνα 19:** Ανάλυση της έκφρασης του lncRNA *CENTRIC2* μεταξύ σταδίων προόδου του γαστρικού καρκίνου (Stage\_I – Stage\_IV) και παρακείμενου φυσιολογικού ιστού γαστρικού καρκίνου (Normal Paracancerous). Τα δείγματα που υπάγονται στην κατηγορία Stage\_NA δεν δύνανται να ταξινομηθούν με αξιοπιστία σε ένα από τα προηγούμενα στάδια.

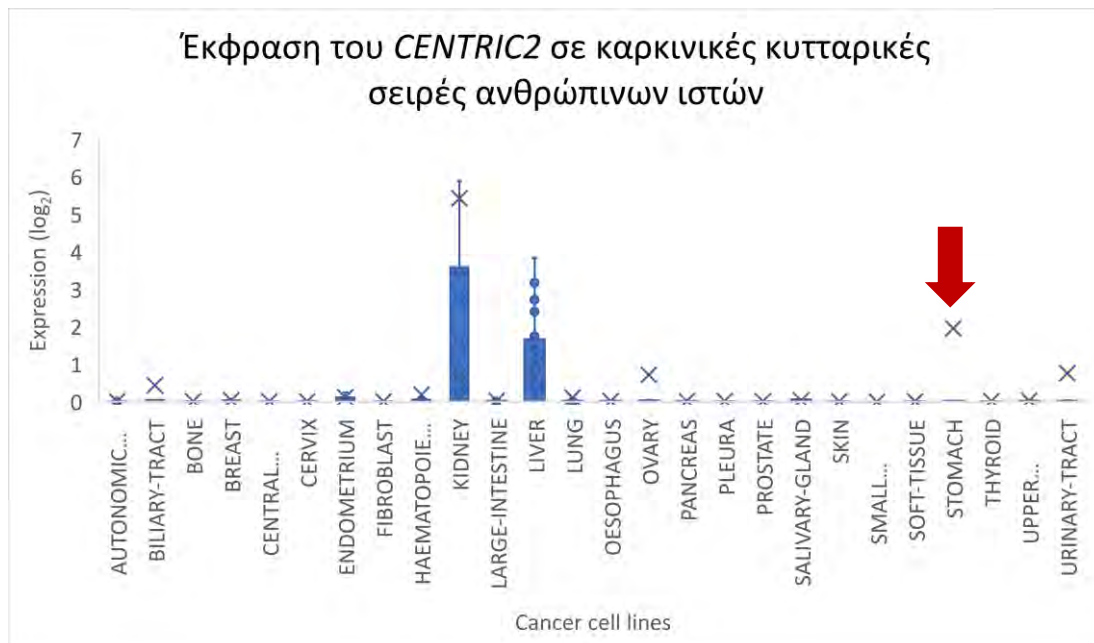
Ακολούθησε ανάλυση δεδομένων RNAseq και για δεδομένα έκφρασης του *CENTRIC2* από στάδια προκαρκινικών αλλοιώσεων γαστρικού καρκίνου που αποτελούν προοίμιο του εντερικού τύπου γαστρικού αδενοκαρκινώματος. Αξιοποιήθηκαν δεδομένα για κάθε ένα από τα στάδια προκαρκινικών αλλοιώσεων (γαστρίτιδας, ατροφίας, εκτεταμένης ατροφίας και εντερικής μεταπλασίας), δεδομένα έκφρασης παρακείμενου φυσιολογικού ιστού γαστρικού καρκίνου, αλλά και δεδομένα έκφρασης από οργανοειδή στομάχου, τα οποία αποτελούν μικρογραφίες ανθρώπινων οργάνων, ανεπτυγμένες μέσω χρήσης καρκινικών βλαστοκυττάρων. Τα αποτελέσματα της ανάλυσης παρατίθενται παρακάτω (**Εικόνα 20**).



**Εικόνα 20:** Ανάλυση της έκφρασης του lncRNA *CENTRIC2* μεταξύ σταδίων προκαρκινικών αλλοιώσεων εντερικού υποτύπου γαστρικού αδενοκαρκινώματος. Διακρίνονται τα στάδια της γαστρίτιδας (*GAST*), της ατροφίας (*AT*), της εκτεταμένης ατροφίας (*EAT*), της εντερικής μεταπλασίας (*IM*), καθώς και δεδομένα έκφρασης από δείγματα παρακείμενου φυσιολογικού ιστού (*NORM PARACANCEROUS*) και οργανοειδών (*ORG*).

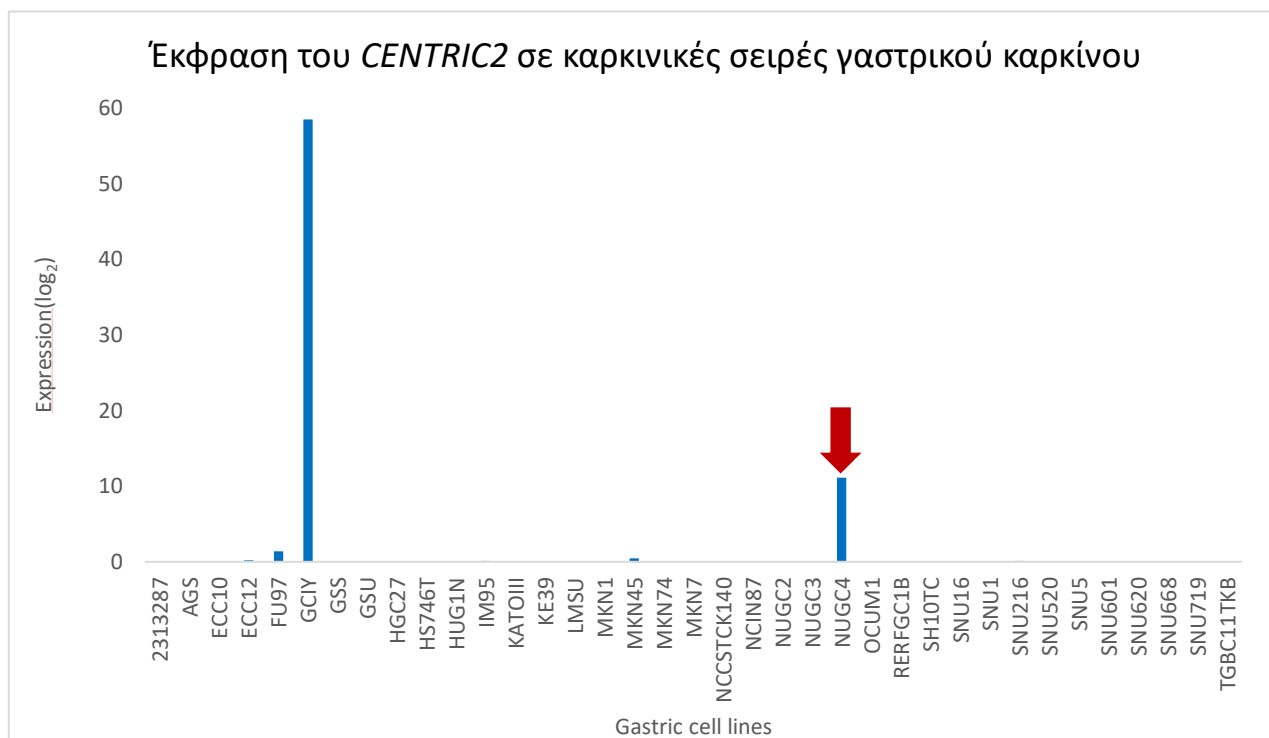
#### 4.1.2.2 Έκφραση του *CENTRIC2* σε δείγματα κυτταρικών σειρών

Επόμενο βήμα ήταν η ανάλυση της έκφρασης του lncRNA *CENTRIC2* σε κυτταρικές σειρές. Αρχικά, αξιοποιήθηκαν δεδομένα έκφρασης από 1020 καρκινικές κυτταρικές σειρές όλων των διαθέσιμων ανθρώπινων ιστών (**Εικόνα 21**).



**Εικόνα 21:** Ανάλυση της έκφρασης του lncRNA *CENTRIC2* σε κυτταρικές σειρές ανθρώπινων φυσιολογικών ιστών. Παρατηρείται υψηλή έκφραση στους νεφρούς, στο ήπαρ, στο στομάχι, στις ωοθήκες και στην χολική και στην ουροποιητική οδό.

Στα πλαίσια της παρούσας διπλωματικής εργασίας, η ανάλυση επικεντρώθηκε σε κυτταρικές σειρές στομάχου. Πιο συγκεκριμένα, αντλήθηκαν και αναλύθηκαν δεδομένα από ένα σύνολο 37 καρκινικών κυτταρικών σειρών στομάχου, τα οποία παρατίθενται παρακάτω (Εικόνα 22).



**Εικόνα 22:** Ανάλυση της έκφρασης του lncRNA CENTRIC2 σε καρκινικές κυτταρικές σειρές στομάχου. Παρατηρείται αυξημένη έκφραση στις κυτταρικές σειρές FU97, GCIY και NUGC4.

#### 4.1.2.3 Ανάλυση συνταινίας του CENTRIC2 στον ποντικό

Η βιοπληροφορική ανάλυση ολοκληρώθηκε με την διεξαγωγή ανάλυσης συνταινίας, προκειμένου να προσδιοριστεί η ύπαρξη πιθανού ορθόλογου γονιδίου σε άλλον οργανισμό. Για την ανάλυση επιλέχθηκε το γονιδίωμα του ποντικού (mm9), συνήθους οργανισμού μοντέλου για πειράματα μελέτης καρκίνου, σε αντιπαραβολή με την έκδοση (hg19) του ανθρώπινου γονιδιώματος. Τα αποτελέσματα της ανάλυσης συνταινίας παρατίθενται παρακάτω (Εικόνα 23).



**Εικόνα 23:** Ανάλυση συνταινίας για το lncRNA CENTRIC2. Απεικονίζονται οι συντεταγμένες του γονιδίου στο ανθρώπινο γονιδίωμα, σε αντιπαραβολή με το γονιδίωμα του ποντικού. Προκύπτει ότι δεν υφίσταται συντήρηση μεταξύ των γονιδιωμάτων των δύο οργανισμών στην συγκεκριμένη περιοχή, υποδηλώνοντας ειδοειδικότητα του lncRNA στον άνθρωπο.



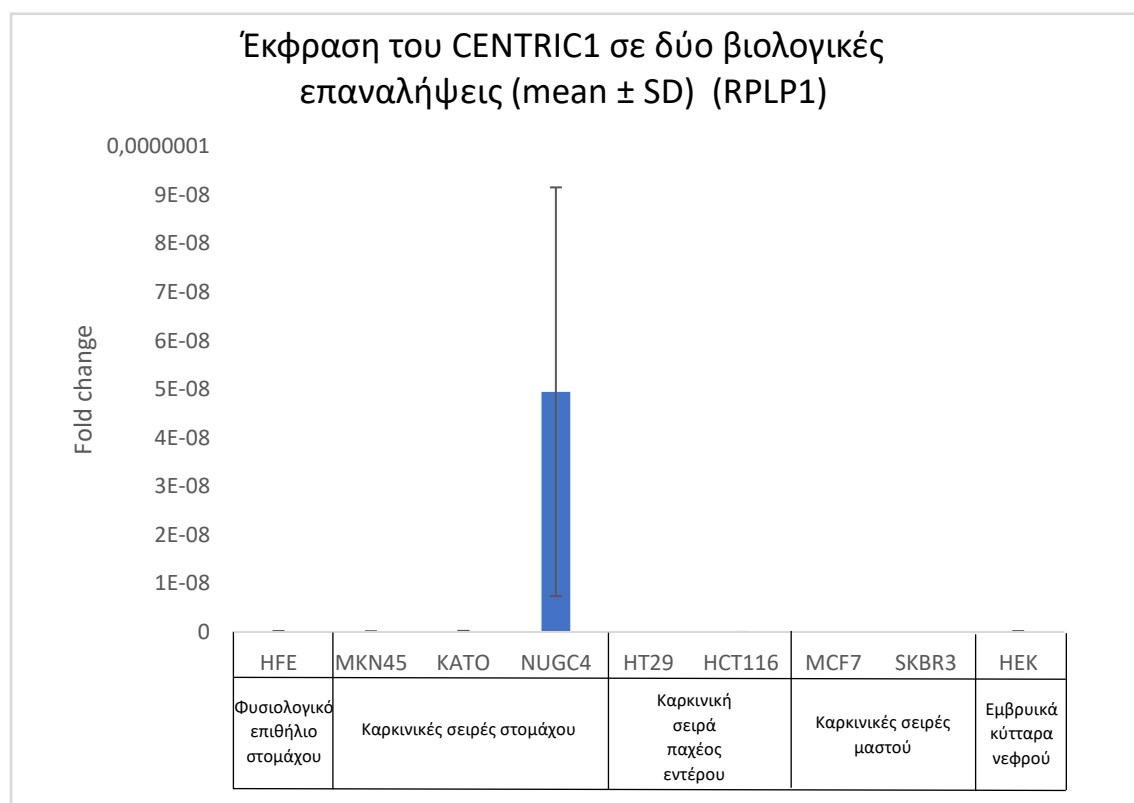
## 4.2 Πειραματικά αποτελέσματα

Το πειραματικό μέρος της παρούσας διπλωματικής εργασίας χωρίστηκε σε δύο μέρη. Το πρώτο πειραματικό μέρος αφορά τον χαρακτηρισμό των lncRNA CENTRIC1 και CENTRIC2 ως προς την εξειδίκευση της έκφρασής τους, τον υποκυτταρικό τους εντοπισμό και την μεταγραφική τους σταθερότητα. Το δεύτερο πειραματικό μέρος αφορούσε τη διαδικασία κατασκευής επιθυμητών εργαλείων γενετικής μηχανικής με στόχο την επιτυχημένη καταστολή των δύο lncRNAs σε κύτταρα καρκινικών σειρών στομάχου, καθώς και τον έλεγχο του φαινοτύπου των κυττάρων σε συνθήκες καταστολής.

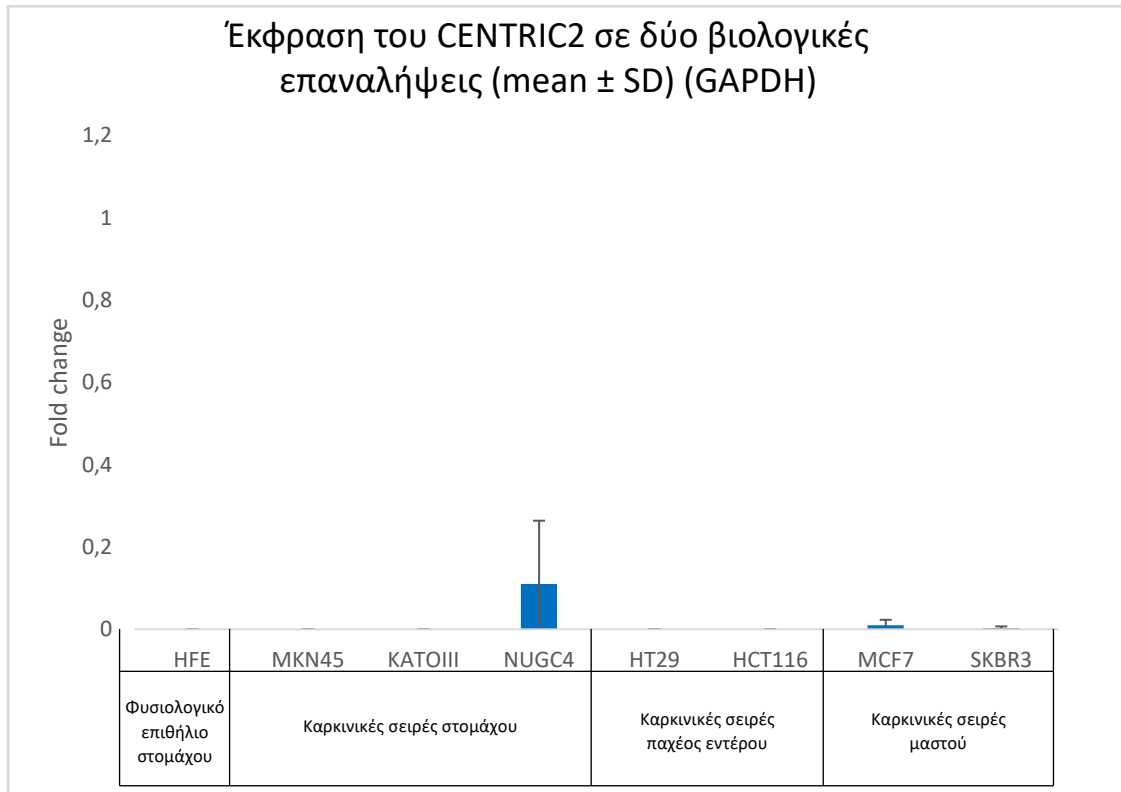
### 4.2.1. 1<sup>ο</sup> Πειραματικό μέρος

#### 4.2.1.1 Έκφραση του lncRNA CENTRIC2 σε ένα σύνολο διαθέσιμων κυτταρικών σειρών

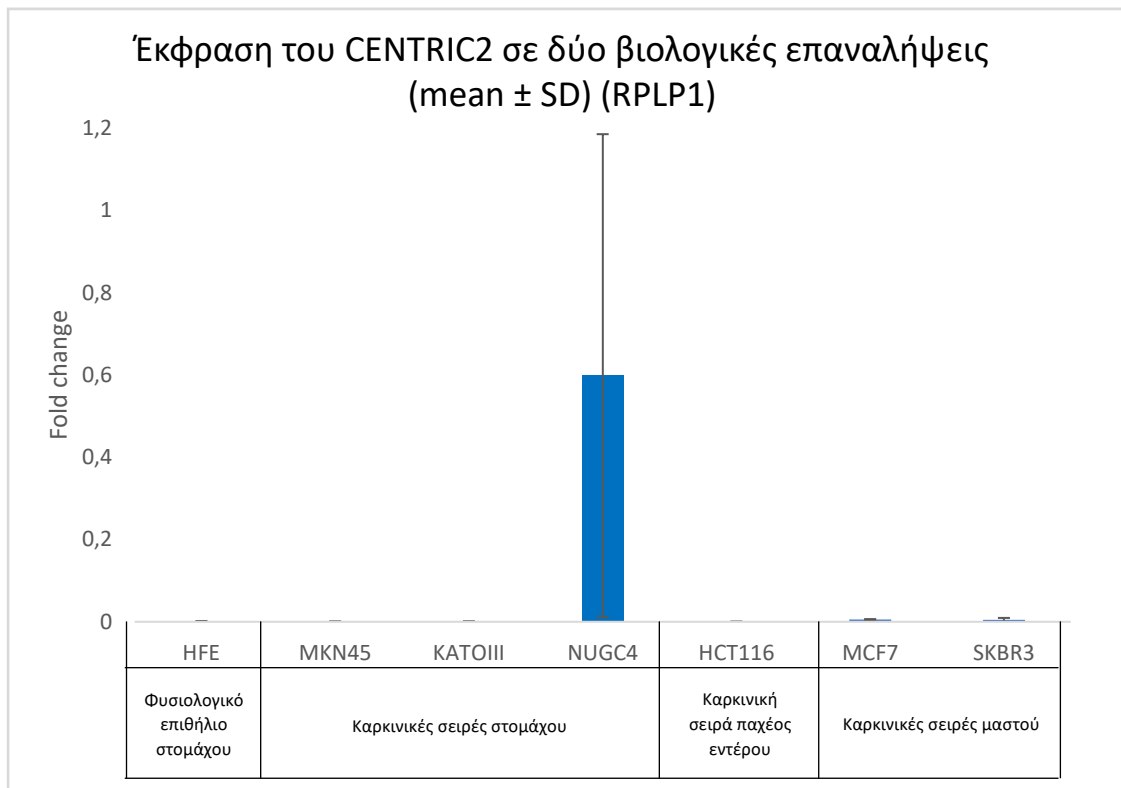
Για το συγκεκριμένο πείραμα πραγματοποιήθηκε απομόνωση RNA, σύνθεση cDNA και qPCR σε διαθέσιμες καρκινικές κυτταρικές σειρές του εργαστηρίου, προκειμένου να εξεταστεί ο πιθανός ρόλος των lncRNAs σε διάφορους τύπους καρκίνου. Ενδεικτικά, αξιοποιήθηκαν τρεις καρκινικές σειρές στομάχου (MKN45, KATOIII, NUGC4), δύο καρκινικές σειρές παχέος εντέρου (HT29, HCT116) δύο καρκινικές σειρές μαστού (MCF7, SKBR3) μία κυτταρική σειρά φυσιολογικού επιθηλίου στομάχου (HFE), καθώς και εμβρυικά κύτταρα νεφρού (HEK). Το σύνολο των παραπάνω κυτταρικών σειρών επιλέχθηκε με σκοπό τον έλεγχο εξειδίκευσης της έκφρασης σε κύτταρα γαστρικού καρκίνου. Για την κανονικοποίηση της έκφρασης των lncRNAs, χρησιμοποιήθηκαν εκκινητές και για δύο γονίδια με ιδιοσυστάτα υψηλά επίπεδα έκφρασης των GAPDH και RPLP1. Οι δύο αναλύσεις απεικονίζουν το εύρος αλλαγής της έκφρασης και αφορούν τον μέσο όρο δύο βιολογικών επαναλήψεων για την έκφραση του lncRNA σε σχέση με το εκάστοτε γονίδιο κανονικοποίησης (**Εικόνα 24A,B,Γ**).



**Εικόνα 24A:** Ανάλυση της έκφρασης του lncRNA CENTRIC1 σε ένα σύνολο καρκινικών κυτταρικών σειρών. Η έκφραση φαίνεται να είναι σημαντικά υψηλότερη στην κυτταρική σειρά NUGC4.



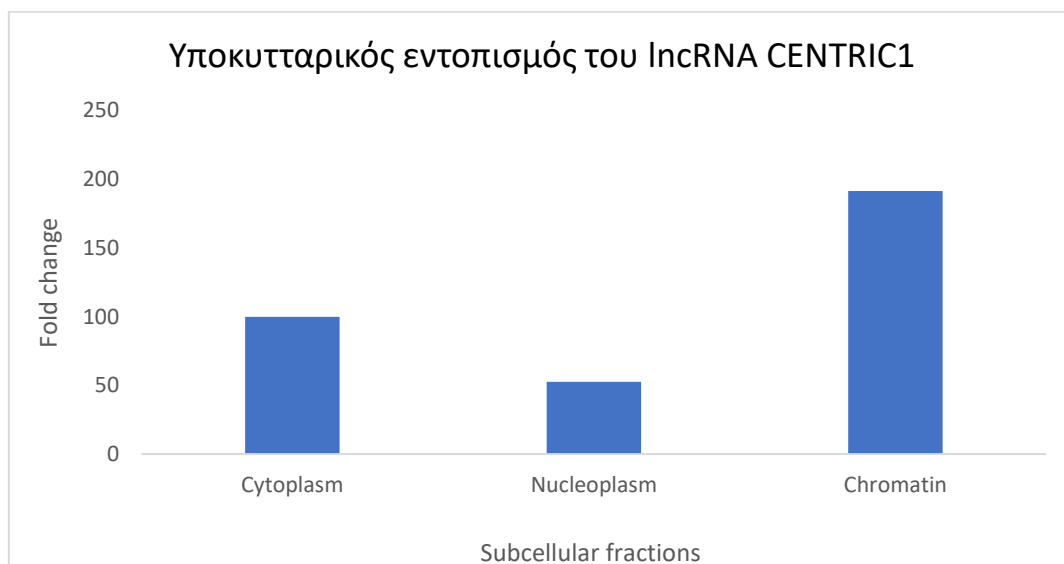
**Εικόνα 24B:** Ανάλυση της έκφρασης του lncRNA CENTRIC2 σε ένα σύνολο καρκινικών κυτταρικών σειρών. Η έκφραση φαίνεται να είναι σημαντικά υψηλότερη στην κυτταρική σειρά NUGC4.



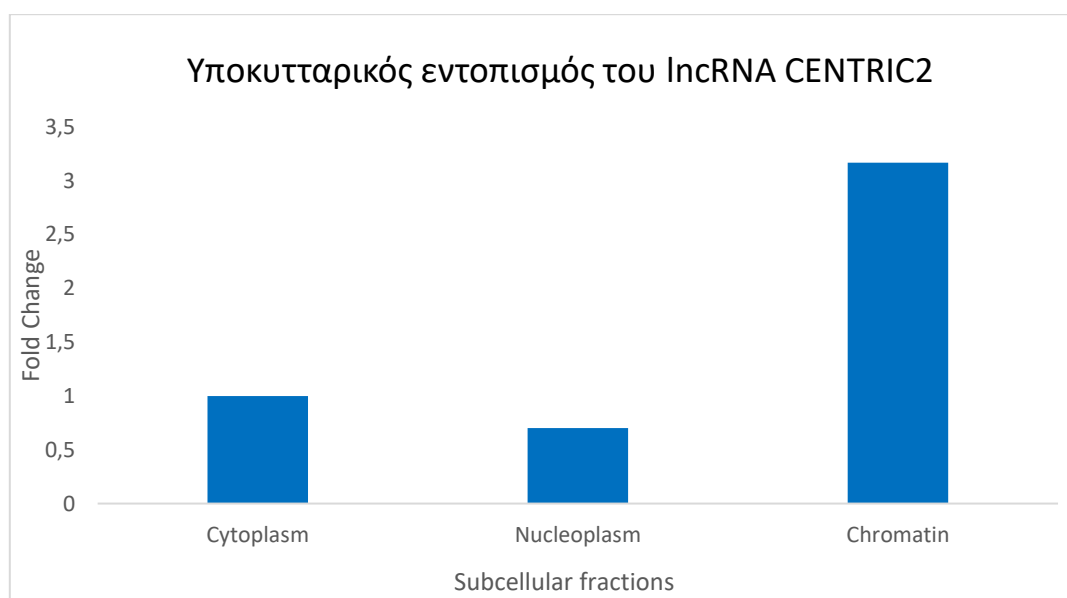
**Εικόνα 24Γ:** Ανάλυση της έκφρασης του lncRNA CENTRIC2 σε ένα σύνολο καρκινικών κυτταρικών σειρών. Η έκφραση φαίνεται να είναι σημαντικά υψηλότερη και σχεδόν εξειδικευμένη στην κυτταρική σειρά NUGC4.

#### 4.2.1.2 Υποκυτταρικός εντοπισμός των lncRNA CENTRIC1 και CENTRIC2

Προκειμένου να προσδιοριστεί το υποκυτταρικό κλάσμα στο οποίο εντοπίζονται τα lncRNAs CENTRIC1 και CENTRIC2 στο κύτταρο, διεξήχθησαν πειράματα υποκυτταρικής κλασμάτωσης σε κύτταρα NUGC4. Συλλέχθηκαν συνολικά τρία κλάσματα, κυτταρόπλασμα, πυρηνόπλασμα και χρωματίνη, στα οποία διενεργήθηκε απομόνωση RNA, σύνθεση cDNA και qPCR. Ανάλογα με την συγκέντρωσή τους στο εκάστοτε κλάσμα μπορεί να εξαχθεί ένα συμπέρασμα για τον εντοπισμό τους και κατά συνέπεια την πιθανή λειτουργία τους στο κύτταρο. Τα σχετικά αποτελέσματα παρατίθενται παρακάτω (**Εικόνα 25A,B**).



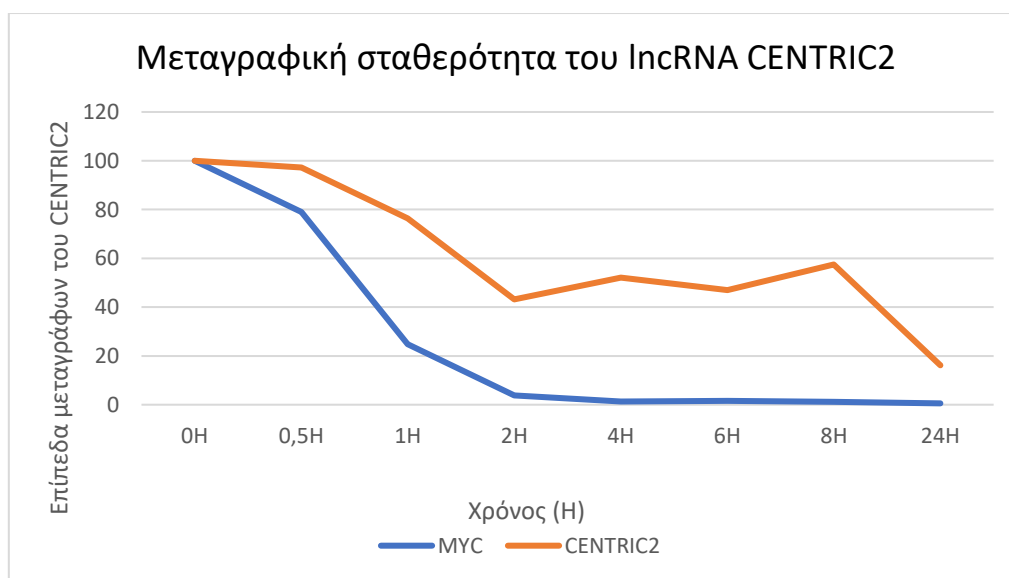
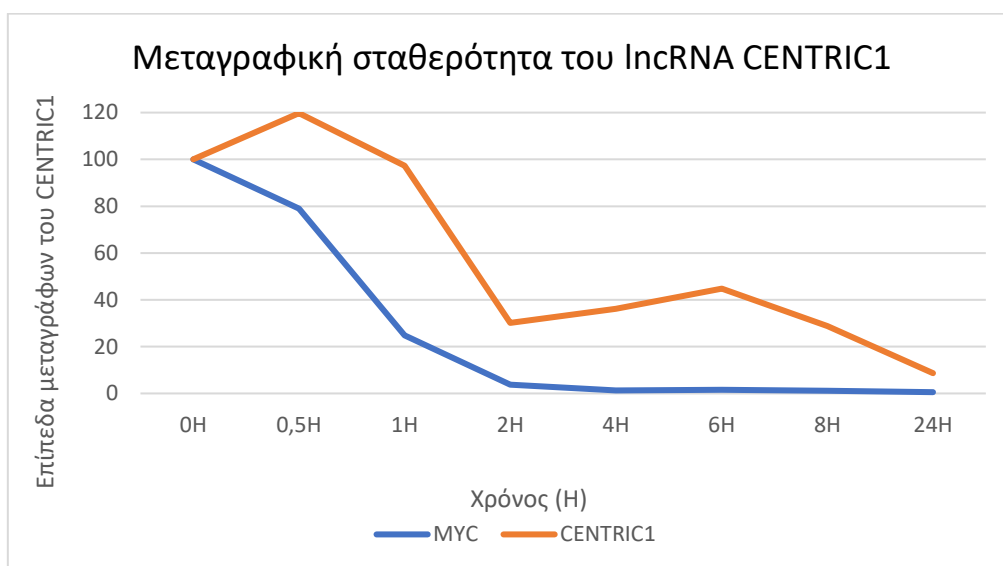
**Εικόνα 25A:** Υποκυτταρικός εντοπισμός του lncRNA CENTRIC1, έπειτα από υποκυτταρική κλασμάτωση καρκινικών γαστρικών κυττάρων NUGC4. Το CENTRIC1 φαίνεται να είναι ένα ως επί το πλείστον χρωματινικό lncRNA.



**Εικόνα 25B:** Υποκυτταρικός εντοπισμός του lncRNA CENTRIC2, έπειτα από υποκυτταρική κλασμάτωση καρκινικών γαστρικών κυττάρων NUGC4. Το CENTRIC2 φαίνεται να είναι ένα ως επί το πλείστον χρωματινικό lncRNA.

#### 4.2.1.3 Μεταγραφική σταθερότητα των lncRNA CENTRIC1 και CENTRIC2

Προκειμένου να επιτευχθεί ο βέλτιστος δυνατός χαρακτηρισμός των lncRNAs CENTRIC1 και CENTRIC2, είναι απαραίτητος ο προσδιορισμός του χρόνου υποδιπλασιασμού των δύο μεταγράφων στο κύτταρο. Η σταθερότητα ενός μορίου RNA καθορίζει σε μεγάλο ποσοστό τις λειτουργίες του και κατά συνέπεια την σημασία του ρόλου του για την επιτέλεση ζωτικών λειτουργιών. Πιο συγκεκριμένα, πραγματοποιήθηκε έκθεση καρκινικών κυττάρων της καρκινικής γαστρικής σειράς NUGC4 σε ακτινομυκίνη, η οποία καταστέλλει την διαδικασία της μεταγραφής και οδηγεί στην αποικοδόμηση των μεταγράφων εντός του κυττάρου. Ακολούθησε απομόνωση RNA, σύνθεση cDNA και qPCR σε διαφορετικά timepoints (0, 0,5, 1, 2, 4, 6, 8, 24 ώρες), ώστε να προσδιοριστούν τα επίπεδα των δύο μεταγράφων. Τα αποτελέσματα απεικονίζονται στα παρακάτω διαγράμματα (Εικόνα 26A,B).



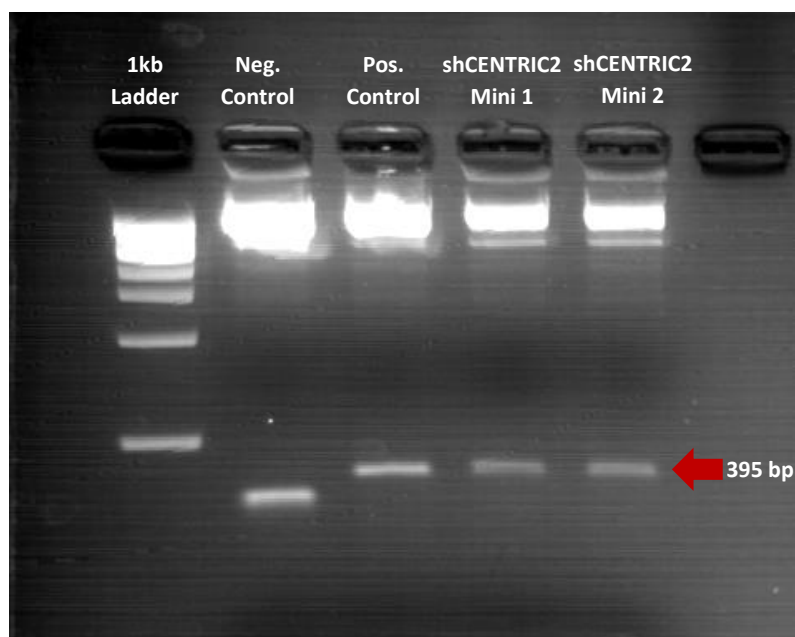
**Εικόνα 26A,B:** Υπολογισμός της μεταγραφικής σταθερότητας των lncRNA CENTRIC1 και CENTRIC2 σε κύτταρα NUGC4. Απεικονίζονται τα επίπεδα των μεταγράφων σε συνάρτηση με ορισμένα χρονικά σημεία σε μια χρονική περίοδο (0, 0,5, 1, 2, 4, 6, 8, 24 ώρες). Ως δείγμα 0H ορίζεται μια ποσότητα κυττάρων στα οποία δεν έχει χορηγηθεί ακτινομυκίνη και σηματοδοτεί την έναρξη του πειράματος. Ο χρόνος υποδιπλασιασμού εντοπίζεται στις 2-4 ώρες.

Όπως υποδεικνύεται από τη βιοπληροφορική ανάλυση και τα πειραματικά δεδομένα, τα lncRNA CENTRIC1 και CENTRIC2 παρουσιάζουν ιδιαίτερο ερευνητικό ενδιαφέρον. Υπό το πρίσμα της μελέτης τους, διεξήχθη μια σειρά πειραματικών διαδικασιών για την καταστολή τους και την ακόλουθη μελέτη του ρόλου τους σε κύτταρα NUGC4. Τα αποτελέσματα παρουσιάζονται στο δεύτερο πειραματικό μέρος.

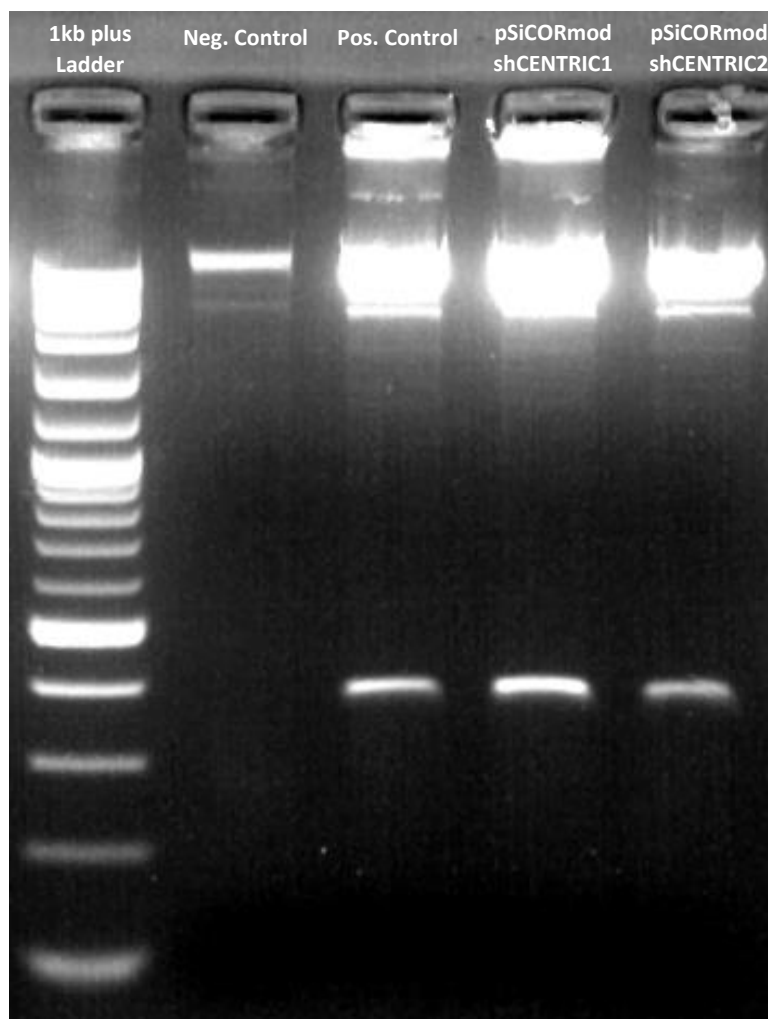
#### 4.2.2. 2<sup>ο</sup> Πειραματικό μέρος

##### 4.2.2.1 Κατασκευή εργαλείων γενετικής μηχανικής για την καταστολή των lncRNAs CENTRIC1 και CENTRIC2

Βασική προϋπόθεση για την πλήρη αποσαφήνιση των ρόλων και των λειτουργιών των CENTRIC1 και CENTRIC2 είναι ο έλεγχος του φαινοτύπου των κυττάρων σε συνθήκες καταστολής της έκφρασής τους. Προκειμένου να εξασφαλισθεί αυτή η συνθήκη, κατασκευάστηκε ένα σύστημα παροδικής καταστολής. Το σύστημα αυτό περιλαμβάνει ανασυνδυασμό του φορέα κλωνοποίησης pSiCORMod με ένθεμα που κωδικοποιεί για ένα μόριο shRNA. Από το shRNA προκύπτει ένα μόριο siRNA, το οποίο αξιοποιεί τον μηχανισμό RNAi του κυττάρου, στοχεύοντας το μετάγραφο στόχο και οδηγώντας το σε αποικοδόμηση. Στα πλαίσια της παρούσας διπλωματικής εργασίας κατασκευάστηκαν δύο ανασυνδυασμένοι φορείς κλωνοποίησης, ένας για την καταστολή του lncRNA CENTRIC1 και ένας για την καταστολή του lncRNA CENTRIC2. Η επιτυχημένη ενσωμάτωση των ενθεμάτων στον φορέα κλωνοποίησης επιβεβαιώνεται με διαγνωστική πέψη (Εικόνα 27A,B) και αλληλούχηση.



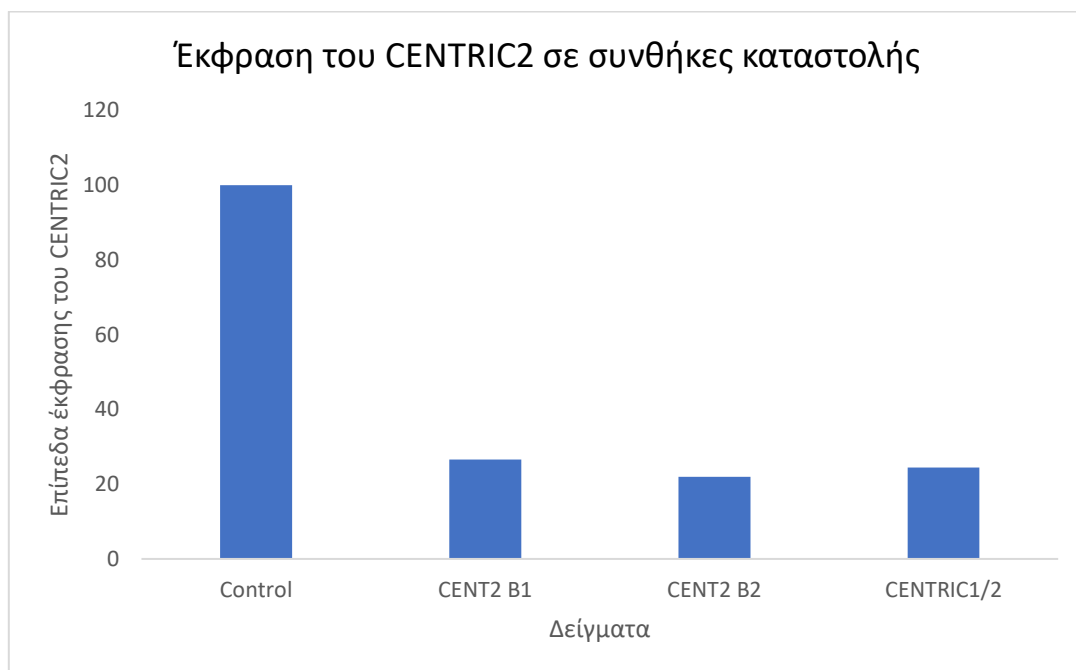
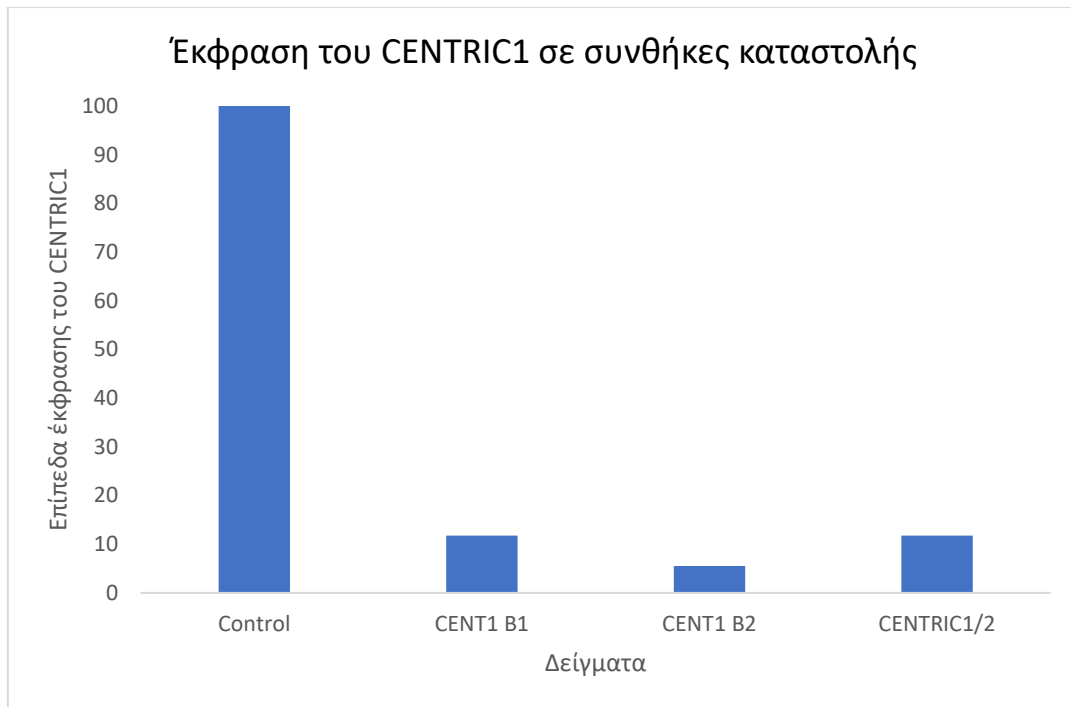
**Εικόνα 27A:** Ηλεκτροφόρηση προϊόντων διαγνωστικής πέψης του φορέα κλωνοποίησης με ένθεμα που στοχεύει στην καταστολή του lncRNA CENTRIC2 σε gel αγαρόζης 2%. Διακρίνονται ο δείκτης μοριακού βάρους (1 kb Ladder), το αρνητικό (Neg. Control) και το θετικό (Pos. Control) δείγμα ελέγχου, καθώς και ο έλεγχος (screening) δύο αποικιών βακτηρίων XL1Blue για τον επιτυχημένο ανασυνδυασμό, έπειτα από απομόνωση πλασμιδίου μικρής κλίμακας (shCENTRIC2 Mini 1, shCENTRIC2 Mini 2). Τα θετικά δείγματα αντιστοιχούν σε μοριακό βάρος 395 bp.



**Εικόνα 27B:** Ηλεκτροφόρηση προϊόντων διαγνωστικής πέψης των φορέων κλωνοποίησης με ενθέματα που στοχεύουν στην καταστολή των lncRNA CENTRIC1 και CENTRIC2 σε gel αγαρόζης 2%. Διακρίνονται ο δείκτης μοριακού βάρους (**1 kb plus Ladder**), το αρνητικό (**Neg. Control**) και το θετικό (**Pos. Control**) δείγμα ελέγχου, καθώς και δύο επιτυχώς ανασυνδυασμένοι φορείς κλωνοποίησης (**pSiCORmod shCENTRIC1**, **pSiCORmod shCENTRIC2**). Ως αρνητικό δείγμα ελέγχου χρησιμοποιήθηκε φορέας pSiCORmod ο οποίος επανακυκλοποιήθηκε κατά την αντίδραση λιγάσης μετά την πέψη (*self ligated vector*). Ως θετικό δείγμα ελέγχου χρησιμοποιήθηκε επιβεβαιωμένα ανασυνδυασμένος φορέας με ένθεμα siRNA ολιγονουκλεοτίδιο για την καταστολή του μεταγράφου του μεταγραφικού παράγοντα GATA6.

#### 4.2.2.2. Έλεγχος των επιπέδων έκφρασης των lncRNAs CENTRIC1 και CENTRIC2 σε συνθήκες καταστολής

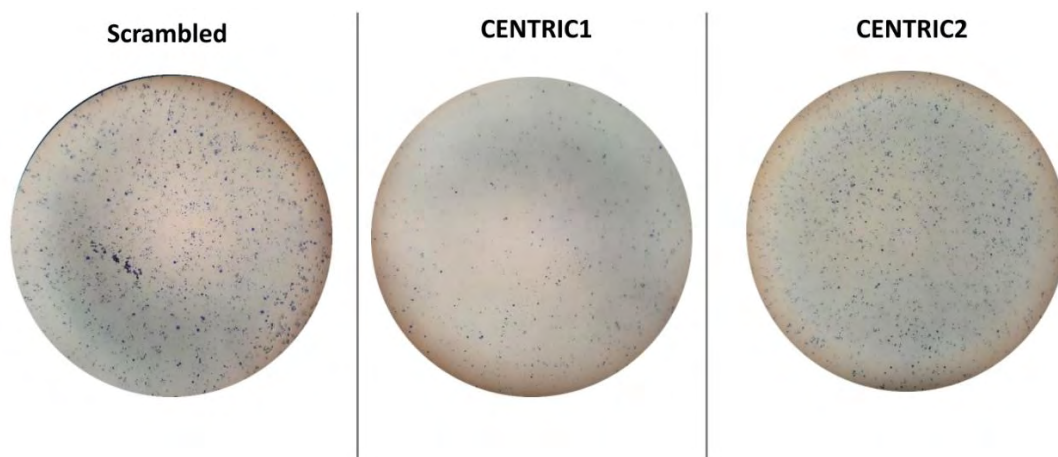
Αφού οι ανασυνδυασμένοι φορείς απομονωθούν σε υψηλή συγκέντρωση, εισάγονται σε ιϊκά σωματίδια λεντιϊών κατά τη συναρμολόγησή τους σε κύτταρα HEK. Τα ιϊκά σωματίδια συλλέγονται από το θρεπτικό μέσο και στη συνέχεια μολύνουν κύτταρα NUGC4, ώστε ο ανασυνδυασμένος φορέας να εισέλθει στο κύτταρο και να καταστείλει την έκφραση του μεταγράφου στόχου. Η επιτυχημένη καταστολή επιβεβαιώνεται με απομόνωση RNA από τα κύτταρα στόχους, σύνθεση cDNA και qPCR (**Εικόνα 28A,B**).



**Εικόνα 28A,B:** Ανάλυση δεδομένων qPCR της έκφρασης σε κύτταρα NUGC4 υπό συνθήκες καταστολής του lncRNA CENTRIC1. Χρησιμοποιήθηκε αρνητικό δείγμα ελέγχου της καταστολής (**Control**) στο οποίο αξιοποιήθηκε φορέας κλωνοποίησης με ένθεμα που δεν στοχεύει στην αποικοδόμηση οποιουδήποτε μεταγράφου στο κύτταρο (*scrambled*). Η καταστολή και ο έλεγχος της έκφρασης πραγματοποιήθηκαν αποκλειστικά για το CENTRIC1 σε δύο βιολογικές επαναλήψεις (**CENT1 B1**, **CENT1 B2**) καθώς και σε πείραμα συνδυαστικής καταστολής με το lncRNA CENTRIC2 (**CENTRIC1/2**). Οι αντίστοιχες αναλύσεις πραγματοποιήθηκαν και για το lncRNA CENTRIC2.

#### 4.2.2.3 Έλεγχος του φαινοτύπου σε επίπεδο κυττάρου υπό συνθήκες καταστολής των lncRNAs CENTRIC1 και CENTRIC2

Το τελικό πείραμα της παρούσας διπλωματικής εργασίας αφορούσε την διεξαγωγή δοκιμασίας ελέγχου σχηματισμού αποικιών (CFA), για την παρατήρηση του φαινοτύπου σε μακροσκοπικό επίπεδο. Η δοκιμασία πραγματοποιήθηκε υπό συνθήκες καταστολής για κάθε lncRNA ξεχωριστά, προκειμένου να ανιχνευθούν πιθανές μεταβολές στο δυναμικό πολλαπλασιασμού των κυττάρων. Τα αποτελέσματα παρατίθενται παρακάτω (**Εικόνα 29**).



**Εικόνα 29:** Δοκιμασία ελέγχου σχηματισμού αποικιών σε κύτταρα NUGC4 υπό συνθήκες καταστολής των lncRNAs CENTRIC1 (κέντρο) και CENTRIC2 (δεξιά). Ως αρνητικό δείγμα ελέγχου (αριστερά) χρησιμοποιήθηκε δείγμα κυττάρων NUGC4 στο οποίο αξιοποιήθηκε φορέας κλωνοποίησης με ένθεμα που δεν στοχεύει στην αποικοδόμηση οποιουδήποτε μεταγράφου στο κύτταρο (Scrambled). Παρατηρήθηκε έντονη αναστολή της κυτταρικής ανάπτυξης κατά την καταστολή του CENTRIC1 ενώ λιγότερη έντονη ήταν η επίδραση της καταστολής του CENTRIC2 σε αυτό το πείραμα.



## 5. ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΣΥΖΗΤΗΣΗ

Ο καρκίνος του στομάχου κατατάσσεται στην τέταρτη θέση όσον αφορά την συχνότητα διάγνωσης, καθώς και στην τρίτη θέση όσον αφορά τη θνησιμότητα στον παγκόσμιο πληθυσμό<sup>312</sup>. Οι επικρατείς σύγχρονες θεραπευτικές προσεγγίσεις περιλαμβάνουν την χειρουργική επέμβαση για αφαίρεση του καρκινικού ιστού, σε συνδυασμό με εκτομή των προσβεβλημένων λεμφαδένων, εξαιτίας της διήθησης των καρκινικών κυττάρων σε αυτούς<sup>313</sup>. Η νόσος χαρακτηρίζεται από ποικίλα φαινόμενα μετάστασης σε ιστούς με αφετηρία τον γαστρικό, όπως τα οστά, οι πνεύμονες και το ήπαρ<sup>314</sup>. Η εξάπλωση των καρκινικών κυττάρων σε παρακείμενους και μη ιστούς καθιστά αναπόφευκτη την δυσμενή αντιμετώπιση της ασθένειας. Ιδιαίτερα σε τελικά στάδια, όπου ο καρκίνος δεν επιδέχεται χειρουργικής επέμβασης, η αποτροπή εξάπλωσης του όγκου και η παράταση του προσδόκιμου ζωής μέσω θεραπευτικής αγωγής κρίνονται ως η βέλτιστη δυνατή λύση<sup>315</sup>. Επιπλέον μειονέκτημα αποτελεί η ελλιπής και μη έγκαιρη διάγνωση της νόσου, η οποία είναι δυνατή αποκλειστικά σε προχωρημένα στάδια, καθώς και η αναποτελεσματική της πρόγνωση, εξαιτίας της υψηλής πιθανότητας υποτροπής. Η ασυμπτωματική φύση του γαστρικού καρκίνου έως τα προχωρημένα στάδια της νόσου επιβάλλει την ανάπτυξη μεθόδων έγκαιρης διάγνωσης, με στόχο την ελάττωση των ποσοστών θνησιμότητας. Κατά συνέπεια, η ανάπτυξη αξιόπιστων και εξειδικευμένων μοριακών βιοδεικτών και θεραπευτικών στόχων κρίνεται απαραίτητη<sup>316</sup>. Η ετερογένεια αποτελεί ένα βασικό γνώρισμα του καρκίνου του στομάχου, σε ιστολογικό, μοριακό αλλά και γενετικό επίπεδο, οδηγώντας στην καθιέρωση των διάφορων υποτύπων γαστρικών αδενοκαρκινωμάτων<sup>69,317</sup>. Σύμφωνα με το πρότυπο μοριακής ταξινόμησης, οι περιπτώσεις γαστρικών αδενοκαρκινωμάτων εντάσσονται σε τέσσερις βασικούς υπότυπους: στον EBV<sup>+</sup>, καθώς και στους υπότυπους που χαρακτηρίζονται από αστάθεια μικροδορυφόρων, γονιδιωματική σταθερότητα ή γονιδιωματική αστάθεια<sup>69</sup>.

Το φαινόμενο της γονιδιωματικής ή χρωμοσωμικής αστάθειας συνιστά τον κύριο μοριακό υπότυπο γαστρικού καρκίνου, αφού εντοπίζεται σε ένα ιδιαίτερα μεγάλο ποσοστό (50%) και συνδέεται στενά με τον εντερικό κατά Lauren ιστολογικό υπότυπο<sup>318</sup>. Χαρακτηρίζεται από υψηλή συχνότητα μεταλλάξεων στο γονιδίωμα ή ακόμα και στο πλαίσιο μιας κυτταρικής γενεαλογίας. Οι μεταλλάξεις αυτές μπορεί να σχετίζονται με μεταβολές στην αλληλουχία των νουκλεϊκών οξέων, χρωμοσωμικές ανακατατάξεις, ανευπλοειδίες, καθώς και παραλλαγές στον αριθμό των γονιδιακών αντιγράφων<sup>69</sup>. Τα Copy Number Variations αποτελούν πρωταρχικό γνώρισμα της γονιδιωματικής αστάθειας, μέσω της μεταβολής της δομής των χρωμοσωμάτων και κατ' επέκταση του αριθμού των γονιδιακών αντιγράφων. Περιλαμβάνουν ελλείψεις, ενθέσεις και διπλασιασμούς αλληλουχιών, φαινόμενα τα οποία είναι αποτέλεσμα σφαλμάτων των επιδιορθωτικών μηχανισμών του DNA<sup>319</sup>. Φυσική συνέπεια αποτελεί η μεταβολή της γονιδιακής δόσης σε λειτουργικά, δομικά ή ρυθμιστικά στοιχεία του γονιδιώματος. Αύξηση ή μείωση της γονιδιακής δόσης σε σχέση με την φυσιολογική διπλοειδή μορφή, ενδέχεται να μεταβάλλει την λειτουργία των γονιδίων στόχων και να οδηγήσει σε ποικίλες ασθένειες, μεταξύ αυτών και σε καρκίνο, μέσω της ενεργοποίησης ογκογόνων μονοπατιών σηματοδότησης<sup>131</sup>. Έχουν καταγραφεί περιπτώσεις μεταγραφικών παραγόντων και άλλων πρωτεϊνών, των οποίων τα γονίδια παρουσιάζουν ενίσχυση στον γενετικό τους τόπο. Η πλειοψηφία των CNVs αντιστοιχούν σε γονίδια τα οποία κωδικοποιούν πρωτεΐνες, κυρίως μεταγραφικούς παράγοντες<sup>320</sup>. Ωστόσο, υπάρχουν περιπτώσεις καρκίνων στους οποίους εντοπίζονται CNVs σε γονίδια που κωδικοποιούν για lncRNAs. Χαρακτηριστικά παραδείγματα είναι αυτά του καρκίνου του στομάχου, του ήπατος, των ωοθηκών, καθώς και του καρκίνου του παχέος εντέρου, στα οποία διεξήχθησαν αναλύσεις δεδομένων έκφρασης και αναπτύχθηκαν εργαλεία υπολογισμού της συσχέτισης

μεταξύ CNVs και lncRNAs. Εντοπίζονται παραλλαγές που αφορούν τόσο αύξηση, όσο και ελάττωση του αριθμού των αντιγράφων, οι οποίες είναι πιθανόν να απορρυθμίζουν την έκφραση των lncRNAs και να τα οδηγούν στο να παρεκκλίνουν από τις φυσιολογικές τους λειτουργίες<sup>321-325</sup>.

Στην παρούσα διπλωματική εργασία, κύριος στόχος ήταν η ταυτοποίηση αντίστοιχων lncRNAs, των οποίων η έκφραση ελέγχεται από CNVs στον καρκίνο του στομάχου, καθώς και η ακόλουθη μελέτη και αποσαφήνιση του ρόλου τους με τη βοήθεια βιοπληροφορικών και πειραματικών μεθόδων. Σύμφωνα με την βιοπληροφορική ανάλυση και το κωδικοποιητικό τους δυναμικό, τα lncRNAs CENTRIC1 και CENTRIC2 φαίνεται να πληρούν τα παραπάνω κριτήρια, καθώς η έκφρασή τους παρουσιάζει θετική συσχέτιση με την αύξηση του αριθμού των αντιγράφων στον γενετικό τους τόπο. Ιδιαίτερα εντυπωσιακό είναι επίσης το γεγονός ότι τα εν λόγω lncRNAs εντοπίζονται σε γενετικούς τόπους που δεν περιέχουν κανένα άλλο γονίδιο, καθιστώντας τα μοναδικά όσον αφορά την αλληλεπίδρασή τους με CNVs. Στην περίπτωση του CENTRIC2, δεν παρατηρήθηκε ομόζυγο έλλειμμα (deep deletion) στον αντίστοιχο γενετικό τόπο, υποδεικνύοντας ότι το συγκεκριμένο φαινόμενο είναι πολύ σπάνιο για τον συγκεκριμένο τύπο καρκίνου. Τα αισθητά υψηλότερα επίπεδα έκφρασης των δύο lncRNA σε καρκινικές κυτταρικές σειρές, σε σχέση με την ελάχιστη έκφρασή τους σε φυσιολογικούς ενήλικους ιστούς, τα καθιστούν πιθανούς εξειδικευμένους διαγνωστικούς δείκτες γαστρικού καρκίνου. Η παραπάνω υπόθεση ενισχύεται από το γεγονός ότι τα επίπεδα έκφρασης μεταξύ φυσιολογικών και καρκινικών κυττάρων δεν παρουσιάζουν ανάλογη μεταβολή σε κανένα άλλο είδος ανθρώπινου ιστού. Τα δεδομένα ανάλυσης της έκφρασης μεταξύ καρκινικού και παρακείμενου φυσιολογικού ιστού από βιοψίες πληθυσμών παρουσιάζουν αυξημένη έκφραση στον όγκο, επαληθεύοντας την συμβολή των CENTRIC1 και CENTRIC2 στην ογκογένεση. Ο παρακαρκινικός φυσιολογικός ιστός χρησιμοποιείται αντί φυσιολογικού γαστρικού ιστού ως δείγμα ελέγχου, καθώς προσεγγίζει στο μέγιστο δυνατό βαθμό τον καρκινικό φαινότυπο, ωστόσο δεν εμφανίζει ιστολογικά ευρήματα, αυξάνοντας με αυτό τον τρόπο την αξιοπιστία της ανάλυσης. Τα αποτελέσματα από την ανάλυση της έκφρασης μεταξύ σταδίων προόδου της νόσου υποδηλώνουν ότι τα επίπεδα των δύο lncRNAs παραμένουν σταθερά υψηλά για τα στάδια I-III, ενώ παρουσιάζουν μια σχετική περαιτέρω αύξηση στο στάδιο IV. Κάτι τέτοιο μπορεί να εξηγηθεί σε συνδυασμό με τις πληροφορίες που εξάγονται από την ανάλυση της έκφρασης των CENTRIC1 και CENTRIC2 σε καρκινικές σειρές στομάχου. Τα συγκεκριμένα lncRNAs παρουσιάζουν αυξημένη έκφραση σε καρκινικές σειρές οι οποίες εμφανίζουν μετάσταση σε ασκίτες ή παραγαστρικούς λεμφαδένες, όπως η σειρά NUGC4. Τα στάδια I-III χαρακτηρίζονται κατά κανόνα από παρόμοια επίπεδα μετάστασης σε έναν μικρό αριθμό λεμφαδένων. Το στάδιο IV αντιστοιχεί σε προχωρημένη μετάσταση του γαστρικού καρκίνου σε απομακρυσμένους ιστούς και όργανα του σώματος. Επομένως, η αυξημένη έκφραση των δύο γονιδίων στο στάδιο IV, είναι πιθανή λογική αιτία της εκτεταμένης ενεργοποίησης του μεταστατικού δυναμικού της νόσου. Όσον αφορά την ανάλυση της έκφρασης των CENTRIC1 και CENTRIC2 σε στάδια προκαρκινικών αλλοιώσεων εντερικού υποτύπου γαστρικού καρκίνου, δεν παρατηρείται κάποια σημαντική διαφορά. Αυτό το γεγονός πιθανώς οφείλεται στον περιορισμένο αριθμό διαθέσιμων δειγμάτων για κάθε στάδιο προκαρκινικής αλλοίωσης. Ωστόσο, αυτό που αξίζει να σημειωθεί είναι το γεγονός ότι τα επίπεδα έκφρασης των δύο lncRNAs σε δείγματα όγκου είναι τουλάχιστον δύο τάξεις μεγέθους υψηλότερα από αυτά των προκαρκινικών αλλοιώσεων, ενισχύοντας με αυτό τον τρόπο τις υποθέσεις για τους ρόλους τους ως πιθανά ογκογονίδια. Τέλος, η μηδενική συντήρηση που προκύπτει από την ανάλυση συνταινίας πιθανώς να οφείλεται σε κάποιο έλλειμμα μεταξύ των γονιδιωμάτων των δύο οργανισμών.

Παρόλα αυτά, είναι πιθανή η συντήρηση σε επίπεδο λειτουργίας ενός τέτοιου μορίου, καθώς η λειτουργία του εκάστοτε lncRNA εξαρτάται από την δευτεροταγή του δομή και όχι από την αλληλουχία του. Επομένως, δεν αποκλείεται η ύπαρξη ενός ορθόλογου γονιδίου του *CENTRIC2* στον ποντικό με εξειδικευμένη λειτουργία στον γαστρικό καρκίνο, το οποίο ωστόσο δεν εκφράζεται στο ανθρώπινο γονιδίωμα.

Στο πλαίσιο των πειραματικών αποτελεσμάτων, η ανάλυση της qPCR σε καρκινικές κυτταρικές σειρές επιβεβαιώνει την ιδιαίτερα υψηλή έκφραση των *CENTRIC1* και *CENTRIC2* αποκλειστικά στην καρκινική σειρά NUGC4. Επιπλέον, τα μηδαμινά επίπεδα έκφρασης σε άλλες γαστρικές καρκινικές σειρές, σε καρκινικές σειρές της γαστρεντερικής οδού και άλλων ανθρώπινων ιστών, υποδηλώνουν ένα εξειδικευμένο πρότυπο έκφρασης τόσο στον γαστρικό καρκίνο, αλλά και πιο συγκεκριμένα στο στάδιο της μετάστασης σε παραγαστρικούς λεμφαδένες. Το πείραμα της υποκυτταρικής κλασμάτωσης οδηγεί στο συμπέρασμα ότι τα *CENTRIC1* και *CENTRIC2* είναι δύο αμιγώς χρωματινικά, μη κωδικοποιητικά μετάγραφα. Ταυτόχρονα, τα δύο lncRNAs φαίνεται να διαθέτουν έναν χρόνο υποδιπλασιασμού γύρω στις 2-4 ώρες, όπως προκύπτει με αξιοπιστία από το πείραμα μεταγραφικής σταθερότητας, καθώς ο χρόνος υποδιπλασιασμού του γονιδίου δείκτη MYC ανέρχεται στο γνωστό διάστημα των 30 λεπτών. Όπως προκύπτει, τα συγκεκριμένα μετάγραφα φαίνεται να παρουσιάζουν υψηλότερο χρόνο υποδιπλασιασμού, καθώς και μεγαλύτερο αριθμό αντιγράφων εντός του κυττάρου, αφού ανέρχονται στους 25-26 κύκλους κατά τη διάρκεια μιας qPCR, σε σύγκριση με τα περισσότερα χαρακτηρισμένα lncRNAs. Τα παραπάνω στοιχεία δικαιολογούν πιθανές *trans-acting* λειτουργίες, καθώς τα επαρκή επίπεδα λειτουργικών μεταγράφων και ο υψηλός χρόνος παραμονής στο κύτταρο είναι απαραίτητες προϋποθέσεις για την προσέγγιση ενός απομακρυσμένου στόχου. Ευνοϊκά συμπεράσματα δύνανται να εξαχθούν και για τον εντοπισμό τους στη χρωματίνη, ως πιθανά μόρια αγκυροβόλησης (scaffolds), με σκοπό τη στρατολόγηση παραγόντων που προσδένονται στο DNA, όπως μεταγραφικοί παράγοντες, RNA πολυμεράσες ή μεθυλοτρανσφεράσες.

Τόσο οι βιοπληροφορικές αναλύσεις, όσο και τα πειραματικά δεδομένα, ανέδειξαν τα lncRNAs *CENTRIC1* και *CENTRIC2* ως μόρια με υψηλό ερευνητικό ενδιαφέρον και οδήγησαν στην ανάπτυξη των κατάλληλων εργαλείων γενετικής μηχανικής που αφορούν την παροδική καταστολή τους σε κύτταρα NUGC4. Τα δεδομένα της ηλεκτροφόρησης από την διαγνωστική πέψη των ανασυνδυασμένων pSiCORmod, σε συνδυασμό με τα δεδομένα αλληλούχησης, επιβεβαίωσαν την επιτυχημένη ένθεση των shRNA ολιγονουκλεοτιδίων στον φορέα κλωνοποίησης. Οι δύο ανασυνδυασμένοι pSiCORmod αξιοποιήθηκαν σε πείραμα διαμόλυνσης κυττάρων NUGC4 με τη βοήθεια λεντιών, ώστε να καταστείλουν τα δύο μετάγραφα. Τα αποτελέσματα της qPCR για τα *CENTRIC1* και *CENTRIC2* υπέδειξαν μια καταστολή της τάξης του 90% και 80% αντίστοιχα, τόσο μεμονωμένα όσο και σε συνθήκες ταυτόχρονης καταστολής, επιβεβαιώνοντας τον επιτυχημένο σχεδιασμό των siRNA και κατά συνέπεια την κατασκευή πλήρως λειτουργικών εργαλείων καταστολής των δύο lncRNAs. Ως τελικό πείραμα χαρακτηρισμού, τα αποτελέσματα του Colony Formation Assay για τα δύο μετάγραφα ήταν επίσης ενθαρρυντικά. Πιο συγκεκριμένα, η καταστολή του *CENTRIC1* προκαλεί εμφανή ελάττωση του πολλαπλασιαστικού δυναμικού των κυττάρων (growth defect). Τα αποτελέσματα για το *CENTRIC2* δεν είναι ξεκάθαρα όσον αφορά το πολλαπλασιαστικό δυναμικό, ωστόσο θα διερευνηθούν μελλοντικά άλλοι φαινότυποι, όπως για παράδειγμα το μεταστατικό δυναμικό των κυττάρων ή η ικανότητα κυτταρικής προσκόλλησης και η κινητικότητά τους.

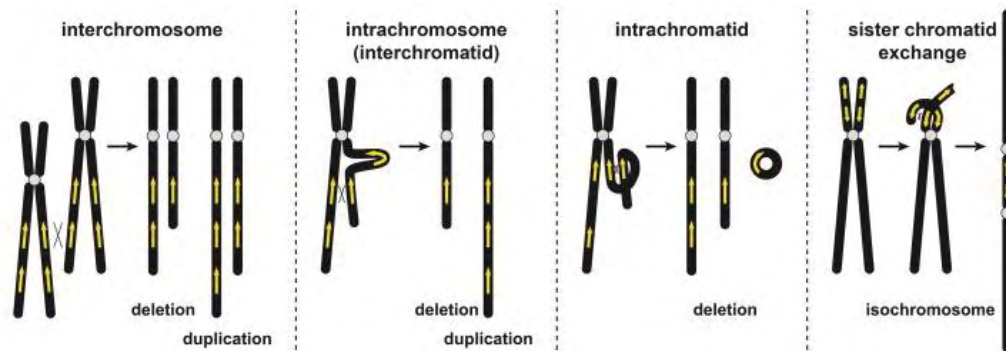
Πρόσθετες ερευνητικές προσεγγίσεις για την περαιτέρω μελέτη του ρόλου των CENTRIC1 και CENTRIC2 περιλαμβάνουν την γενετική τροποποίηση περισσότερων κυτταρικών σειρών γαστρικού καρκίνου (π.χ. AGS) για την ανάπτυξη ενός πληρέστερου προφίλ έκφρασης των γονιδίων στον καρκίνο του στομάχου. Η γενετική τροποποίηση δύναται να πραγματοποιηθεί είτε μέσω παροδικής καταστολής σε περίπτωση που ανιχνευθούν υψηλά επίπεδα έκφρασης, είτε μέσω μεταγραφικής ενεργοποίησης μέσω CRISPR activation σε κυτταρικές σειρές με μηδαμινή έκφραση και ευκολότερη δυνατότητα χειρισμού (π.χ. MKN45). Επιπλέον, η κλωνοποίηση του πλήρους μήκους των δύο μεταγράφων σε φορείς κλωνοποίησης θα ήταν χρήσιμη για τη μελέτη τους σε πειράματα ανοσοκατακρήμνισης RNA (RNA Immunoprecipitation) ή RNA pull-down. Με αυτόν τον τρόπο, θα καταστεί δυνατή η ανίχνευση πιθανών αλληλεπιδράσεων των lncRNAs με πρωτεΐνες και ο ακόλουθος χαρακτηρισμός τους ως συστατικά σηματοδοτικών μονοπατιών ογκογόνων διαδικασιών. Τέλος, η αποσαφήνιση πιθανής συνεργιστικής ή ανταγωνιστικής δράσης μεταξύ των δύο γονιδίων παρουσιάζει μεγάλο ενδιαφέρον. Ωστόσο, χρειάζεται να προηγηθεί ενδεδειγμένος χαρακτηρισμός του κάθε lncRNA ξεχωριστά με επιπλέον μεθόδους (π.χ. FACS) καθώς το αποτέλεσμα ενός CFA από δείγμα κυττάρων διπλής καταστολής ενδέχεται να οφείλεται στην καταστολή αποκλειστικά εκ των δύο lncRNAs.

## 6. ΠΑΡΑΡΤΗΜΑ

## 6.1 Μηχανισμοί σχηματισμού των CNVs

Η αρχιτεκτονική του γονιδιώματος, δηλαδή η οργάνωση των επαναλαμβανόμενων και άλλων αλληλουχιών σε επίπεδο προσανατολισμού, μεγέθους, πυκνότητας και κατανομής, είναι ένα σημαντικό χαρακτηριστικό για την κατανόηση και την πρόβλεψη της γενετικής ποικιλομορφίας<sup>326</sup>. Περίπου 50% του ανθρώπινου γονιδιώματος αποτελείται από επαναλαμβανόμενες αλληλουχίες. Σε αυτές ανήκουν μεταθετά στοιχεία (**mobile elements**), όπως τα Alu, τα οποία έχουν την ικανότητα να μετακινούνται από μια περιοχή του γονιδιώματος σε μια άλλη, επαναλήψεις σε κωδικές περιοχές αλλά και σε μη κωδικές, όπως οι περιοχές των κεντρομερών, των τελομερών και των γονιδίων που κωδικοποιούν τα συστατικά των ριβοσωμάτων<sup>327</sup>. Αξιοσημείωτη είναι η περίπτωση των επαναλήψεων μικρού αριθμού αντιγράφων (**Low Copy Repeats - LCRs**) οι οποίες διαμορφώνουν ένα σύνολο ξεχωριστών δομικών παραλλαγών που ονομάζονται τμηματικοί διπλασιασμοί (**Segmental Duplications - SDs**)<sup>104,105,326</sup>. Τα SDs ορίζονται ως τμήματα του γονιδιώματος τα οποία ταυτίζονται μεταξύ τους σε ποσοστό άνω του 90%, όσον αφορά την αλληλουχία τους. Έχουν μήκος συνήθως άνω του 1kb και καταλαμβάνουν το 4-5% του ανθρώπινου γενετικού υλικού, περισσότερο από κάθε άλλο είδος LCR. Αυτό τα αναδεικνύει και ως τις κύριες επαναλήψεις που συμβάλλουν στην δημιουργία γενετικής ποικιλομορφίας<sup>328,329</sup>.

Η πλειοψηφία των LCRs μεγάλου μήκους (άνω των 10kb), δημιουργούν συστοιχίες παράλογων αλληλουχιών, οι οποίες έχουν διαφορετική προέλευση. Ο μοριακός χαρακτηρισμός τέτοιων επαναλαμβανόμενων περιοχών στο ανθρώπινο γονιδίωμα έχει αποκαλύψει μια αρχιτεκτονική μωσαϊκού<sup>330,331</sup>. Η συγκεκριμένη αρχιτεκτονική απαντάται στον τρόπο οργάνωσης τους, δηλαδή σε ιεραρχημένες ομάδες αλληλουχιών με κανονική ή ανεστραμμένη κατεύθυνση. Το παραπάνω χαρακτηριστικό είναι και αυτό που διαφοροποιεί τα LCRs μεγάλου μήκους από τους διπλασιασμούς που προκύπτουν μεταξύ ομόλογων χρωμοσωμάτων ή αδελφών χρωματίδων<sup>332-334</sup> (**Εικόνα 30**). Η κατανομή τους στο γονιδίωμα υποδεικνύει μια αλληλοεπικάλυψη με περιοχές που εμφανίζουν υψηλή συχνότητα ανακατατάξεων, γεγονός που τις συνδέει με έναν μεγάλο αριθμό ασθενειών<sup>329</sup>.



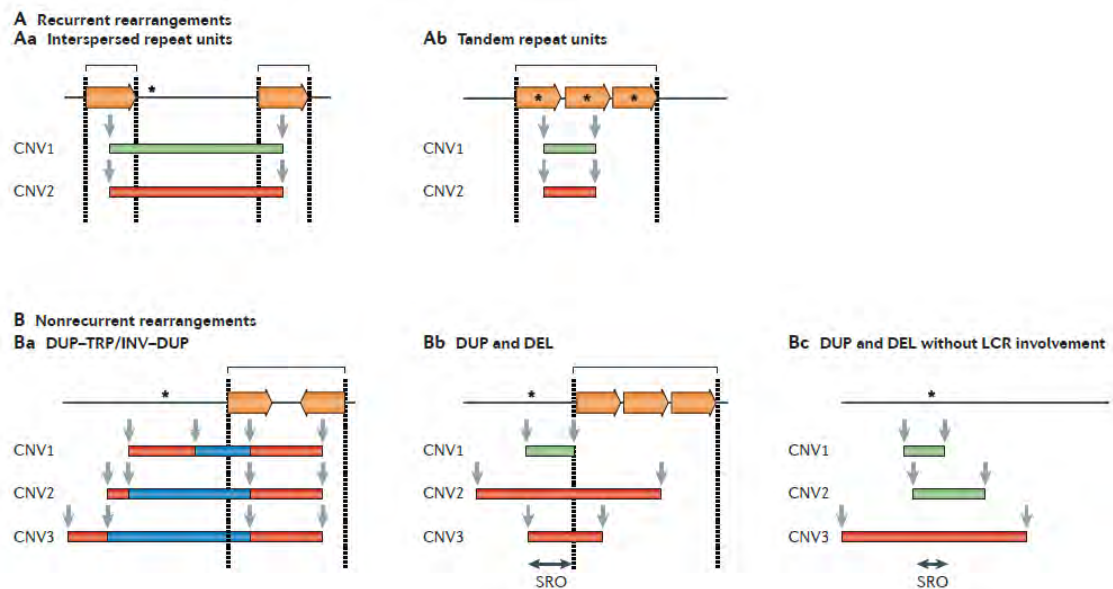
**Εικόνα 30:** Διπλασιασμοί και ελλείψεις που προκύπτουν από αλληλεπίδραση μεταξύ ομόλογων χρωμοσωμάτων ή αδελφών χρωματίδων<sup>335</sup>.

Η αλλαγή στον αριθμό των αντιγράφων προϋποθέτει μια μεταβολή την δομή των χρωμοσωμάτων, κατά κανόνα μέσω της ένωσης δύο αλληλουχιών DNA που έχουν προηγουμένως χωριστεί μεταξύ τους. Η ανάλυση τέτοιων ενώσεων παρέχει σημαντικές πληροφορίες σχετικά με την προέλευση της χρωμοσωμικής δομής<sup>104</sup>.

Κάθε ανακατάταξη που εντοπίζεται στο γονιδίωμα μπορεί να χαρακτηριστεί ως επανεμφανιζόμενη (**recurrent rearrangement**) ή μη επανεμφανιζόμενη (**non-recurrent rearrangement**)<sup>104,105,326,336</sup>. Οι επανεμφανιζόμενες ανακατατάξεις αφορούν τμήματα ενός

γονιδιακού τύπου με σταθερό μέγεθος και περιεχόμενο που επανεμφανίζονται στον πληθυσμό, είναι δηλαδή κοινές μεταξύ μη συγγενικών ατόμων. Τα άκρα αυτών των τμημάτων εντοπίζονται συνήθως στο εσωτερικό εξαιρετικά πανομοιότυπων αλληλουχιών μεγάλου μήκους, συνήθως LCRs<sup>334</sup>. Στο εσωτερικό τους ενδέχεται να περιέχονται γονίδια, των οποίων η λειτουργία πιθανώς επηρεάζεται από το gene dosage effect. Αυτή η κατηγορία ανακατατάξεων ευθύνεται σε μεγάλο ποσοστό για την δημιουργία παθολογικών CNVs<sup>337</sup>. Επανεμφανιζόμενες δομικές παραλλαγές δύνανται επίσης να προκύψουν μεταξύ συνεχόμενων παράλογων γονιδίων και να μεταβάλλουν με αυτό τον τρόπο τον αριθμό των αντιγράφων τους<sup>338</sup>.

Οι μη επανεμφανιζόμενες ανακατατάξεις αφορούν τμήματα ξεχωριστού μεγέθους και περιεχομένου ενός γονιδιακού τύπου που δεν είναι κοινά μεταξύ μη συγγενικών ατόμων. Γονίδια στόχοι μπορούν να ταυτοποιηθούν για την ευαισθησία τους στο gene dosage effect, μέσω της ελάχιστης επικαλυπτόμενης περιοχής (**smallest region of overlap - SRO**) σε ένα σύνολο ασθενών με παρόμοιους κλινικούς φαινότυπους<sup>339</sup>. Ορισμένες κατηγορίες μη επανεμφανιζόμενων ανακατατάξεων παρατηρούνται σε μεγάλο ποσοστό στις γονιδιωματικές διαταραχές. Η οριοθέτησή τους εξυπηρετείται, όπως και στην περίπτωση των επανεμφανιζόμενων ανακατατάξεων, με τη βοήθεια LCRs. Γενετικοί τύποι με υψηλή συχνότητα ανακατατάξεων μπορούν να εμφανίσουν ένα πολύπλοκο πρότυπο δομικών παραλλαγών. Για παράδειγμα, οι παράλογες ανάστροφες επαναλήψεις οδηγούν στον σχηματισμό μιας δομής γνωστής με το όνομα **DUP-TRP/INV-DUP (duplication-inverted triplication-duplication)**<sup>340-344</sup>. Η συγκεκριμένη δομή περιλαμβάνει ένα σύνολο διπλασιασμών και τριπλασιασμών τόσο σε ευθεία, όσο και σε ανάστροφη διάταξη. Οι ανάστροφες επαναλήψεις της συγκεκριμένης δομής αντιστοιχούν στην πλειοψηφία τους σε LCRs αλλά και σε στοιχεία Alu, τα οποία έχουν συσχετιστεί με την εκδήλωση ασθενειών<sup>345</sup> (**Εικόνα 31**).



**Εικόνα 31:** Σχηματική απεικόνιση των επανεμφανιζόμενων και μη επανεμφανιζόμενων γονιδιωματικών ανακατατάξεων<sup>104</sup>.



Οι δύο αυτές γενικές κατηγορίες γονιδιωματικών ανακατατάξεων, παρουσιάζουν μια τάση για δημιουργία περίπλοκων δομικών παραλλαγών, καθώς και εγγενώς διακριτά χαρακτηριστικά που αντικατοπτρίζουν τους υποκείμενους μηχανισμούς σχηματισμού τους.

### Μηχανισμοί σχηματισμού δομικών παραλλαγών

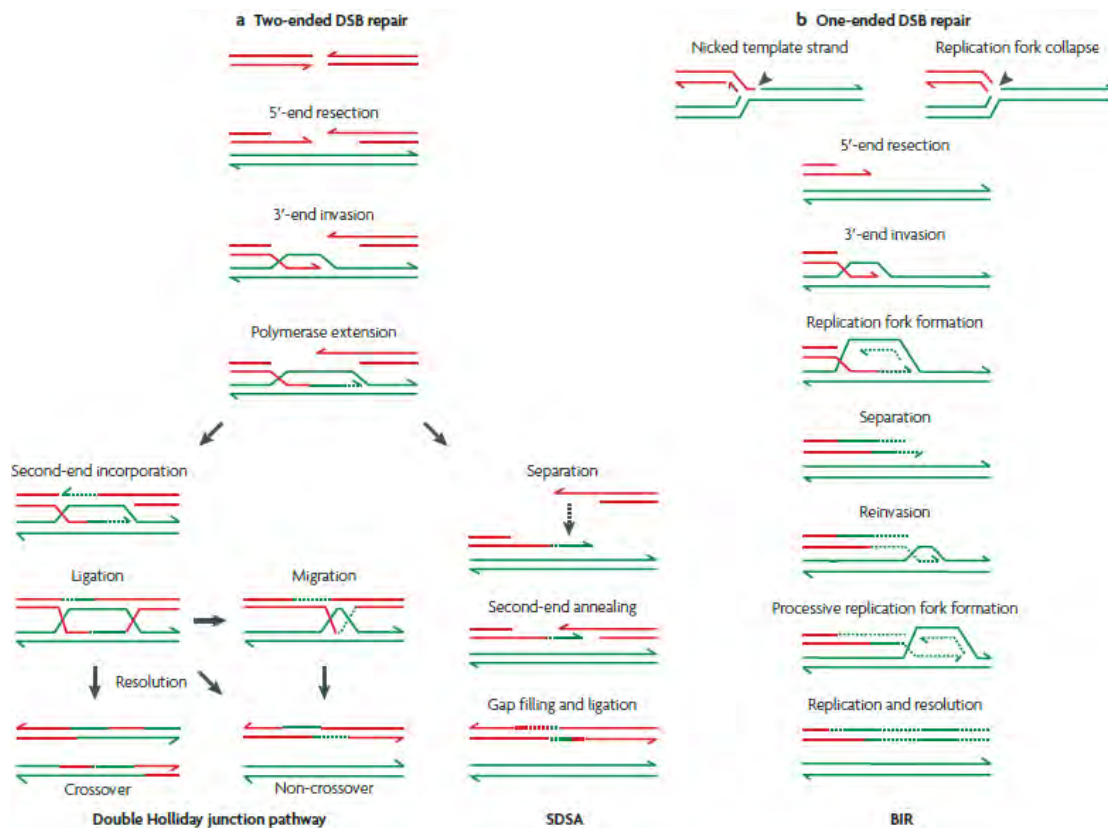
Όπως προαναφέρθηκε, η μεταβολή στον αριθμό των αντιγράφων περιλαμβάνει αλλαγές στην δομή των χρωμοσωμάτων, τέτοιες ώστε προηγουμένως χωριστές χρωμοσωμικές περιοχές να βρίσκονται πλέον κοντά ή μία στην άλλη. Οι μεταβολές αυτές προκύπτουν μέσω δύο βασικών μηχανισμών: του ομόλογου ανασυνδυασμού (**Homologous Recombination - HR**) και του μη ομόλογου ανασυνδυασμού (**Non-Homologous Recombination**)<sup>104</sup>.

Ο ομόλογος ανασυνδυασμός αποτελεί τη βάση διάφορων μηχανισμών επιδιόρθωσης του DNA που προϋποθέτουν την ύπαρξη μιας ομόλογης αλληλουχίας για την επισκευή μιας ελαττωματικής. Ο μηχανισμός απαιτεί συνήθως αλληλουχία μήκους περίπου 300 ζευγών βάσεων στον άνθρωπο καθώς και την ύπαρξη μιας πρωτεΐνης ανταλλαγής κλώνου (**strand exchange protein**)<sup>326,346</sup>. Η αλλαγή στον αριθμό των αντιγράφων, δύναται να προκύψει μέσω ομόλογου ανασυνδυασμού εξαιτίας της παρουσίας LCRs στο γονιδίωμα και όχι όπως θα περίμενε κανείς λόγω της μειωμένης αξιοπιστίας του μηχανισμού. Ωστόσο, κατά τη διαδικασία επιδιόρθωσης, ενδέχεται να χρησιμοποιηθούν ως καλούπι αλληλουχίες ομόλογες με την ελαττωματική, οι οποίες βρίσκονται σε διαφορετική περιοχή στο γονιδίωμα από την επιθυμητή<sup>347</sup>. Το φαινόμενο αυτό ορίζεται ως το μοντέλο της εκτοπικής σύναψης (**ectopic synapsis model**)<sup>326,348</sup>. Ο HR περιλαμβάνει μια ποικιλία διαδικασιών επιδιόρθωσης και είναι υπεύθυνος για τον διαχωρισμό των χρωμοσωμάτων καθώς και για την δημιουργία νέων απλοτύπων κατά τη διάρκεια της μείωσης. Πρωταρχική του λειτουργία είναι η κάλυψη των κενών αλληλουχίας και η επαναφορά των ρήξεων του DNA στην φυσιολογική κατάσταση. Ο πιο καλά χαρακτηρισμένος μηχανισμός του HR είναι αυτός που ενεργοποιείται από την παρουσία δίκλωνων ρήξεων (**DSB - double stranded breaks induced recombination**)<sup>326,349</sup>.

Δύο είναι τα κυριότερα μοντέλα επιδιόρθωσης των DSB<sup>350,351</sup>: Το πρώτο περιλαμβάνει τον σχηματισμό διπλού κόμβου του Holliday (**double Holliday junction**), δηλαδή δύο σημείων στα οποία δύο δίκλινα μόρια DNA ανταλλάσσουν κλώνους μεταξύ τους. Ο συγκεκριμένος μηχανισμός μπορεί να προκαλέσει είτε αμοιβαία μετατόπιση τμημάτων μεταξύ ομόλογων περιοχών (**crossover**), είτε μη αμοιβαία μετατόπιση και μετατροπή του αρχικού αλληλομόρφου (**gene conversion**). Η μετατροπή αυτή φυσιολογικά περιλαμβάνει αλλαγή της αλληλουχίας του κλώνου ενός δίκλωνου μορίου DNA, με στόχο την αποκατάσταση της συμπληρωματικότητας σε περίπτωση σφάλματος<sup>352</sup>. Το δεύτερο μοντέλο είναι η εξαρτώμενη από τη σύνθεση αποδιάταξη των κλώνων του DNA (**Synthesis Dependent Strand Annealing - SDSA**)<sup>353</sup>. Η SDSA φαίνεται να αποτελεί πιθανό μηχανισμό σχηματισμού CNVs, χωρίς ωστόσο να προκαλεί crossovers ή το φαινόμενο της απώλειας ετεροζυγωτίας (**Loss of Heterozygosity - LOH**), δηλαδή της απουσίας ενός αλληλομόρφου μεταξύ δύο χρωμοσωμάτων σε ένα διπλοειδές κύτταρο<sup>326</sup>. Αμοιβαία ανταλλαγή μεταξύ ομόλογων χρωμοσωμάτων μπορεί να οδηγήσει σε LOH εάν οι χρωματίδες που φέρουν ταυτόσημα αλληλόμορφα δεν διαχωριστούν κατά τη μίτωση. Εάν η μετατόπιση λάβει χώρα σε μη ομόλογες περιοχές του ίδιου χρωμοσώματος, αυτό θα έχει ως αποτέλεσμα τον διπλασιασμό ή την έλλειψη αλληλουχιών μέσω του μη αλληλομορφικού ομόλογου ανασυνδυασμού (**Non-Allelic Homologous Recombination - NAHR**)<sup>354</sup> (**Εικόνα 33**). Στην περίπτωση του ανασυνδυασμού μεταξύ αδερφών χρωματίδων, η αμοιβαία μετατόπιση τμημάτων που περιέχουν επαναλήψεις ενδέχεται να οδηγήσει σε ελλείψεις ή αναστροφές<sup>355,356</sup>. Αιτιολογικός παράγοντας των μη αμοιβαίων μετατοπίσεων μπορεί να αποτελέσει η δράση ελικασών και τοποϊσομερασών του DNA, καθώς και το μέγεθος των επαναλήψεων<sup>348,357,358</sup>. Αμοιβαίες μετατοπίσεις πραγματοποιούνται σπανιότερα μεταξύ επαναλήψεων μικρού

μήκους κατά τον ομόλογο ανασυνδυασμό, πιθανώς επειδή υπάρχει μειωμένη ικανότητα σχηματισμού κόμβων του Holliday<sup>359</sup>.

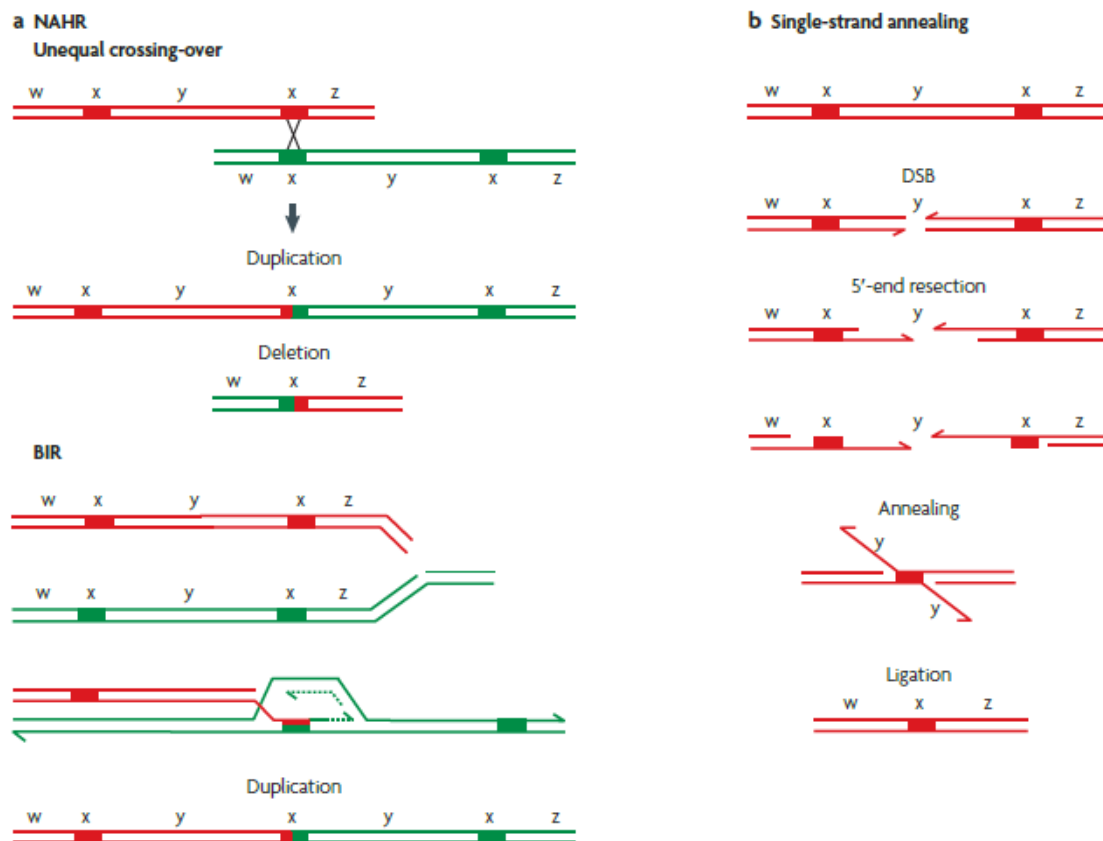
Ο HR δεν χρησιμοποιείται αποκλειστικά για την επιδιόρθωση δίκλωνων ρήξεων, αλλά και για την επιδιόρθωση ελαττωματικών αντιγραφικών διχαλών (**replication forks**). Η διαδικασία που λαμβάνει χώρα ονομάζεται αντιγραφή επαγόμενη από ρήξη (**Break-Induced Replication - BIR**). Υπάρχουν πολλά στοιχεία ότι ο συγκεκριμένος μηχανισμός μπορεί να προκαλέσει δομικές παραλλαγές (έχω citations – mechanisms of change in gene copy number). Φυσιολογικά, η BIR χαρακτηρίζεται από υψηλή πιστότητα και δεν προκαλεί σφάλματα. Ωστόσο, μπορεί να οδηγήσει σε απώλεια ετεροζυγωτίας αν το σημείο της ρήξης βρεθεί κοντά σε μια ομόλογη περιοχή, διαφορετική από αυτή που διαθέτει το ομόλογο χρωμόσωμα του<sup>360-364</sup>. Εάν κατά τη διαδικασία επιδιόρθωσης περιλαμβάνονται ομόλογες αλληλουχίες σε διαφορετικές περιοχές στο γονιδίωμα τότε γεγονότα μετάθεσης, διπλασιασμού ή έλλειψης μπορούν να πραγματοποιηθούν<sup>365</sup> (**Εικόνα 32**).



**Εικόνα 32:** Σχηματική απεικόνιση των μηχανισμών ομόλογου ανασυνδυασμού<sup>326</sup>.

Ελλείψεις μικρής κλίμακας μπορούν να προκύψουν μέσω ενός μηχανισμού επιδιόρθωσης που δρα έναντι διαδοχικά επαναλαμβανόμενων αλληλουχιών. Ο μηχανισμός αυτός, γνωστός ως μηχανισμός αποδιάταξης μονού κλώνου (**Single Strand Annealing - SSA**) ενεργοποιείται όταν κανένα από τα δύο άκρα των DSB δεν βρεθεί κοντά σε ομόλογη αλληλουχία (**Εικόνα 25**). Σε αυτή την περίπτωση, πραγματοποιείται εκτομή των 5' άκρων και ακόλουθη έκθεση του μονόκλωνου 3' άκρου της συμπληρωματικής αλυσίδας. Εάν μέσω αυτής της διαδικασίας προκύψουν φαινόμενα συμπληρωματικότητας σε δύο μονόκλινα τμήματα, ενδέχεται να ακολουθήσει επαναδιάταξη των κλώνων. Αφαίρεση των προεξέχοντων τμημάτων και επακόλουθος σχηματισμός φωσφοδιεστερικού δεσμού ολοκληρώνουν τη διαδικασία επιδιόρθωσης. Αποτέλεσμα αυτής της διαδικασίας θα είναι η αφαίρεση ενός επαναλαμβανόμενου τμήματος<sup>119,366</sup>. Στον άνθρωπο, ο SSA έχει καταγραφεί να δρα σε στοιχεία Alu, τα οποία διαχωρίζονται από μικρό αριθμό ζευγών βάσεων<sup>350,367,368</sup>. Όσο

μεγαλύτερη η αλληλουχία που παρεμβάλλεται μεταξύ των επαναλήψεων τόσο λιγότερο πιθανή είναι η επιτυχία του μηχανισμού. Ο συγκεκριμένος περιορισμός καθιστά τον μηχανισμό ως έναν δευτερεύοντα παίκτη στον σχηματισμό των CNVs (**Εικόνα 33**). Η ποικιλία των τρόπων με τους οποίους προκαλείται δομική παραλλαγή και κατά συνέπεια αλλαγή στον αριθμό των αντιγράφων υποδεικνύει τις πιθανές παρενέργειες του ομόλογου ανασυνδυασμού, παρόλο που παρέχει επιδιορθωτικούς μηχανισμούς ζωτικής σημασίας για το κύτταρο<sup>369</sup>. Προκειμένου να διασφαλίσουν την πιστότητα τους, οι μηχανισμοί ελέγχουν σε υψηλό επίπεδο την επιλογή της πρότυπης αλληλουχίας, καθώς και το επιθυμητό μήκος ομόλογης αλληλουχίας που θα αξιοποιηθεί για την επιδιόρθωση. Εξίσου σημαντική είναι και η μείωση των crossovers στο ελάχιστο δυνατό, κάτι που ωστόσο αποτελεί απαραίτητη διαδικασία στο στάδιο της μείωσης, αυξάνοντας τη συχνότητα εμφάνισης CNV<sup>326</sup>.



**Εικόνα 33:** Σχηματική αναπαράσταση των μηχανισμών ομόλογου ανασυνδυασμού που οδηγούν στον σχηματισμό CNVs<sup>326</sup>.

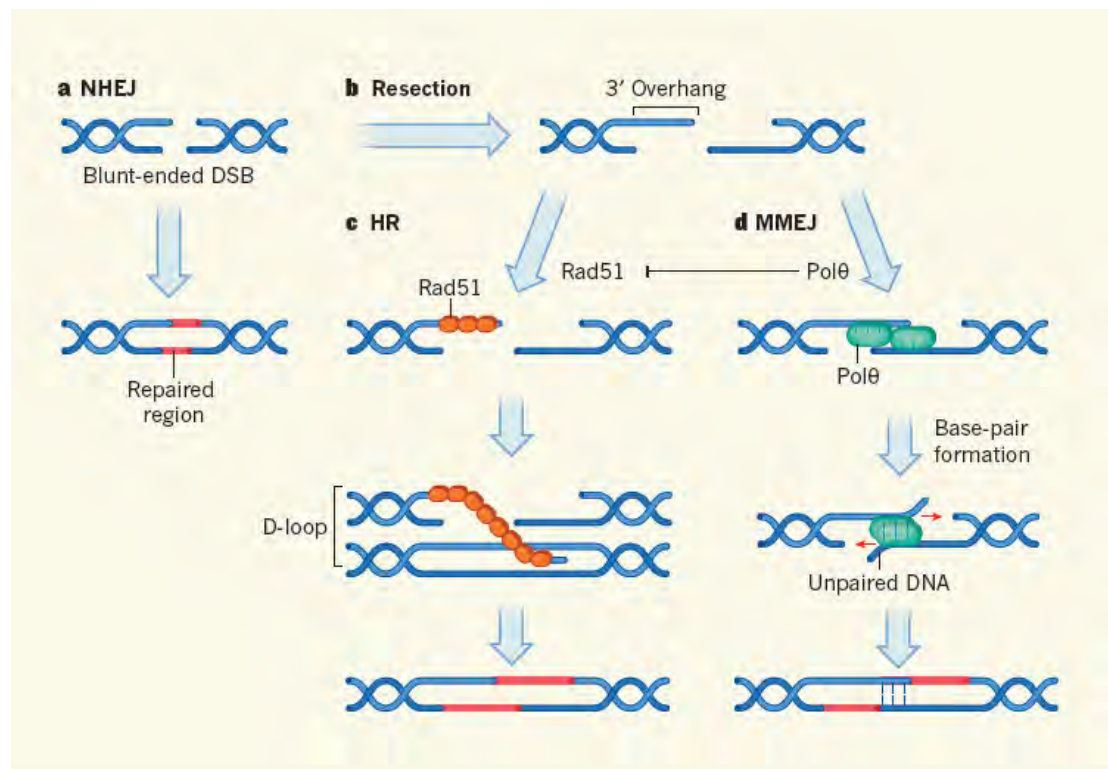
Σε αντίθεση με τον HR, διακρίνονται και μηχανισμοί που χρησιμοποιούν σε περιορισμένο ποσοστό ή και καθόλου τις ομόλογες περιοχές για την επιδιόρθωση του DNA, οδηγώντας σε γενετικές μεταβολές όπως αυτές στον αριθμό των αντιγράφων. Οι συγκεκριμένοι μηχανισμοί χωρίζονται σε δύο ξεχωριστές κατηγορίες: τους αντιγραφικούς (**replicative**) οι οποίοι στηρίζονται σε χαρακτηριστικά της αντιγραφής του DNA και τους μη αντιγραφικούς (**non-replicative**)<sup>370</sup>.

#### **Μη αντιγραφικοί μηχανισμοί επιδιόρθωσης μη ομόλογου ανασυνδυασμού**

**Μη ομόλογη σύνδεση άκρων:** Υπάρχουν δύο μέθοδοι επιδιόρθωσης δίκλωνων ρήξεων που είτε δεν απαιτούν την παρουσία ομόλογων αλληλουχιών είτε αξιοποιούν την ύπαρξη ομολογίας σε επίπεδο μικρού αριθμού βάσεων, φαινόμενο γνωστό και ως μικρο-ομολογία

(**microhomology**)<sup>371</sup>. Οι μέθοδοι αυτές είναι η μη ομόλογη σύνδεση άκρων (**non-homologous end joining - NHEJ**) και η κατευθυνόμενη από την μικρο-ομολογία σύνδεση άκρων (**microhomology-mediated end joining - MMEJ**)<sup>372-374</sup> (Εικόνα 34).

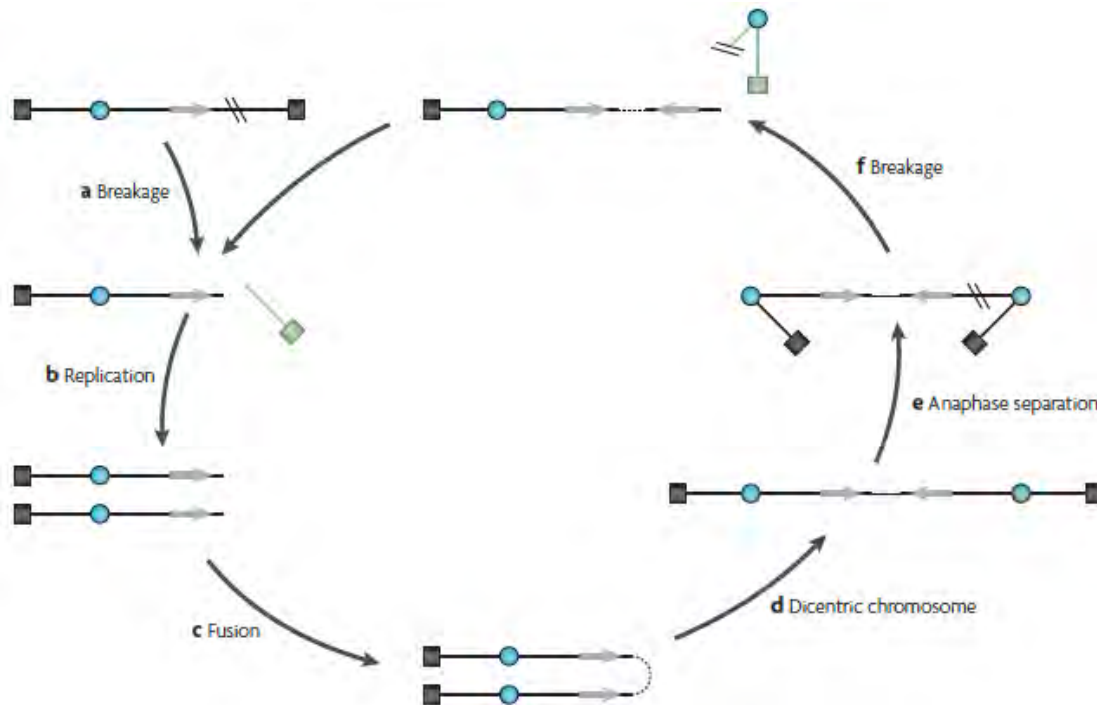
Η μη ομόλογη σύνδεση άκρων χαρακτηρίζεται τόσο από ακρίβεια στην ένωση των άκρων, από ελλείψεις μικρής κλίμακας, 1-4 bp, και σε ορισμένες περιπτώσεις από ενθέσεις ελεύθερου DNA, προερχόμενου συνήθως από μιτοχόνδρια ή ρετροτρανσποζόνια (**retrotransposons**)<sup>375,376</sup>. Κατά την MMEJ ομόλογες περιοχές 5-25 βάσεων επαναδιατάσσονται στα άκρα των DSBs και οδηγούν σε ελλείψεις τμημάτων μεταξύ αυτών, με παρόμοιο τρόπο με τον SSA<sup>326</sup>. Τόσο η NHEJ, όσο και η MMEJ αποτελούν πιθανούς αιτιολογικούς παράγοντες χρωμοσωμικών ανακατατάξεων μέσω της ένωσης μη ομόλογων αλληλουχιών. Κάτι τέτοιο μπορεί να προκύψει σε πολλές περιπτώσεις. Παράδειγμα αποτελούν οι δίκλωνες ρήξεις σε δύο σημεία (**two-ended DSBs**), οι οποίες μπορούν να προκληθούν από εξωγενείς παράγοντες, όπως χημειοθεραπευτικά φάρμακα, ή στην περίπτωση που δύο αντιγραφικές διχάλες υποστούν ταυτόχρονη ρήξη ενός κλώνου τους<sup>377</sup>. Προγραμματισμένες DSBs σε δύο σημεία λαμβάνουν χώρα κυρίως σε κύτταρα του ανοσοποιητικού συστήματος αλλά και κατά την μείωση, όπου ενεργοποιούν μηχανισμούς ομόλογου ανασυνδυασμού<sup>378</sup>.



Εικόνα 34: Μηχανισμοί επιδιόρθωσης των δίκλωνων ρήξεων στο DNA<sup>379</sup>.

**Κύκλος γέφυρας-θραύσης-σύντηξης (Breakage–fusion–bridge cycle - BFB):** Ο συγκεκριμένος μηχανισμός είναι χαρακτηριστικός σε περιπτώσεις γονιδιωματικής αστάθειας και προκαλείται από ενζυμική θραύση των χρωμοσωμάτων ή και αναστολή της σύνθεσης νέων μορίων DNA. Ειδικότερα, συμβαίνει συνήθως έπειτα από την ολοκλήρωση της αντιγραφής ενός χρωμοσώματος το οποίο έχει χάσει τις τελομερικές του περιοχές, εξαιτίας δίκλωνων ρήξεων. Ως αποτέλεσμα, από αυτό το χρωμόσωμα, θα προκύψουν δύο αδελφές χρωματίδες χωρίς τελομερή, οι οποίες θα συντηηθούν δημιουργώντας ένα δικεντρικό χρωμόσωμα, το οποίο αποτελείται από δύο κεντρομερή<sup>380</sup>. Κατά το στάδιο της ανάφασης, τα δύο κεντρομερή

θα χωριστούν σε διαφορετικούς πυρήνες οδηγώντας σε θραύση της ασυνήθιστης αυτής δομής. Η απουσία των τελομερών θα εξακολουθεί να υφίσταται και στις μετέπειτα διαδικασίες αντιγραφής, σχηματίζοντας εκ νέου δικεντρικά χρωμοσώματα και κατ' επέκταση έναν κύκλο γεγονότων (Εικόνα 35). Ο κύκλος τελικά θα αρχίσει να φθίνει σε συχνότητα εμφάνισης όταν το χρωμόσωμα θα αποκτήσει ξανά τελομερικές περιοχές.



Εικόνα 35: Σχηματική απεικόνιση του κύκλου γέφυρας - θραύσης - σύντηξης<sup>326</sup>.

Τυχαίες θραύσεις στα δικεντρικά χρωμοσώματα μπορούν να προκαλέσουν μεγάλης κλίμακας ανάστροφους διπλασιασμούς οι οποίοι ενδέχεται να αυξηθούν σε αριθμό με το πέρασμα των κύκλων. Η συγκεκριμένη διαδικασία έχει συσχετιστεί με γονιδιωματική ενίσχυση στα θηλαστικά και υποστηρίζεται ότι διαδραματίζει καθοριστικό ρόλο στον καρκίνο<sup>381</sup>. Η τυχαία θραύση του χρωμοσώματος αποτελεί εξήγηση για την εμφάνιση ανάστροφων επαναλήψεων σε ανθρώπινα καρκινικά κύτταρα<sup>382</sup>. Γεγονότα μεταβολής που αποδίδονται στον κύκλο BFB μπορούν επίσης να προκληθούν και μέσω οποιασδήποτε διαδικασίας μη ομόλογου ανασυνδυασμού<sup>383</sup>. Όταν μια μετατόπιση πραγματοποιείται σε αντίστροφη διάταξη, δημιουργώντας ένα δικεντρικό χρωμόσωμα, ένα δεύτερο γεγονός πρέπει να λάβει χώρα προκειμένου να εξασφαλιστεί η σταθερότητα του γονιδιώματος<sup>326</sup>. Αυτό το δεύτερο γεγονός μπορεί να αποτελεί συνέχεια της πρώτης μετατόπισης όπως θα περιγραφεί στη συνέχεια.

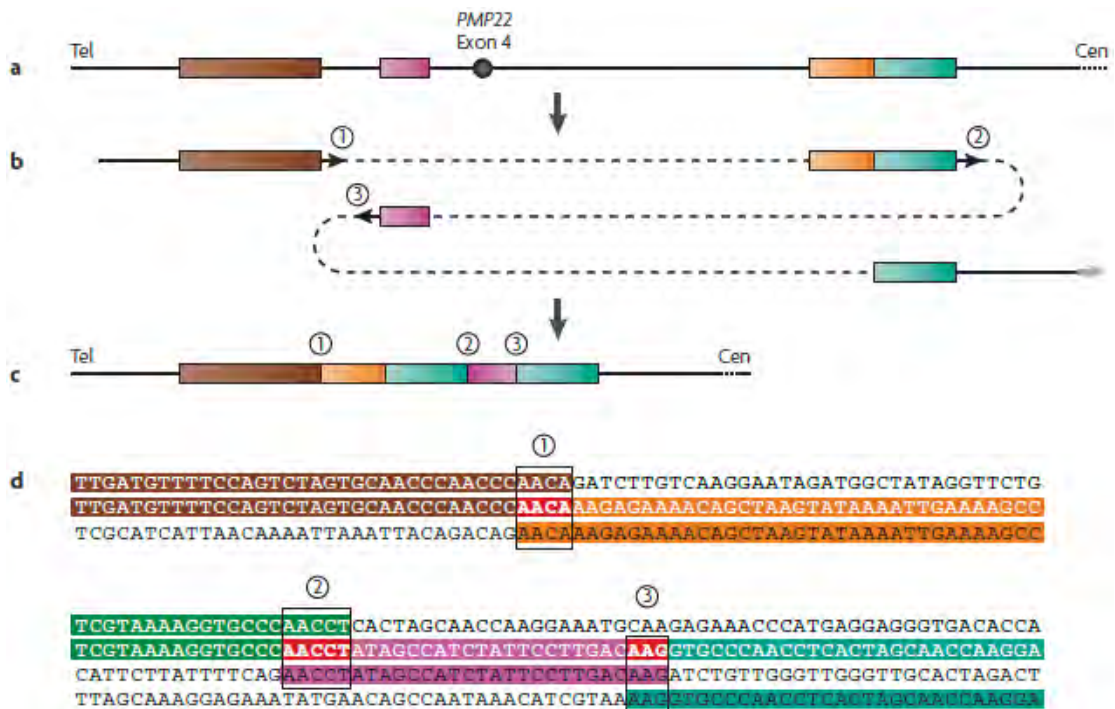
### Αντιγραφικοί μηχανισμοί επιδιόρθωσης μη ομόλογου ανασυνδυασμού

Η εμπλοκή του αντιγραφικού μηχανισμού στον σχηματισμό των CNVs δεν αποτελεί πρόσφατη ανακάλυψη<sup>384-386</sup>. Υπάρχουν συνεχώς αυξανόμενα στοιχεία που υποστηρίζουν ότι διαταραχές στην αντιγραφή μπορούν να ευθύνονται για μεταβολές στον αριθμό των αντιγράφων<sup>383,384,387-390</sup>. Πιο συγκεκριμένα, η αφιδικολίνη, ένας αναστολέας των αντιγραφικών πολυμερασών, προκαλεί τη δημιουργία CNVs σε εύθραυστες χρωμοσωμικές περιοχές κατά μήκος του γονιδιώματος<sup>391-395</sup>. Επιπλέον, είναι γνωστό ότι τα δίκλινα άκρα προκύπτουν από παρεμπόδιση της αντιγραφής<sup>396</sup>. Μελέτες έχουν δείξει ότι το 65% των CNVs

που προκύπτουν από την δράση της αφιδικολίνης, παρουσιάζουν μικρο-ομολογία ή και μηδενική ομολογία στα άκρα τους, γεγονός που απορρίπτει το ενδεχόμενο προέλευσής τους από διαδικασίες ομόλογου ανασυνδυασμού<sup>395</sup>. Όπως προαναφέρθηκε, η παρουσία μικρο-ομολογίας σε μια θέση μη ομόλογου ανασυνδυασμού αποτελεί ορόσημο για μηχανισμούς όπως ο NHEJ. Ωστόσο, έχουν καταγραφεί περιπτώσεις συσχέτισης του φαινομένου με την διαδικασία της αντιγραφής του DNA. Αντιγραφικοί μηχανισμοί έχουν αρχίσει όλο και περισσότερο να θεωρούνται αιτιολογικοί παράγοντες των δομικών παραλλαγών με τη συμμετοχή μικρο-ομόλογων περιοχών. Συνεπώς, παρόλο που και άλλοι μηχανισμοί θα μπορούσαν να εμπλέκονται στην επιδιόρθωση των δίκλωνων ρήξεων, υπάρχουν ενδείξεις ότι αντιγραφικοί μηχανισμοί που στηρίζονται σε μη ομόλογο ανασυνδυασμό θα μπορούσαν να έχουν καθοριστικό ρόλο στην δημιουργία CNVs<sup>326</sup>. Οι εν λόγω μηχανισμοί περιγράφονται παρακάτω (**Εικόνα 37**).

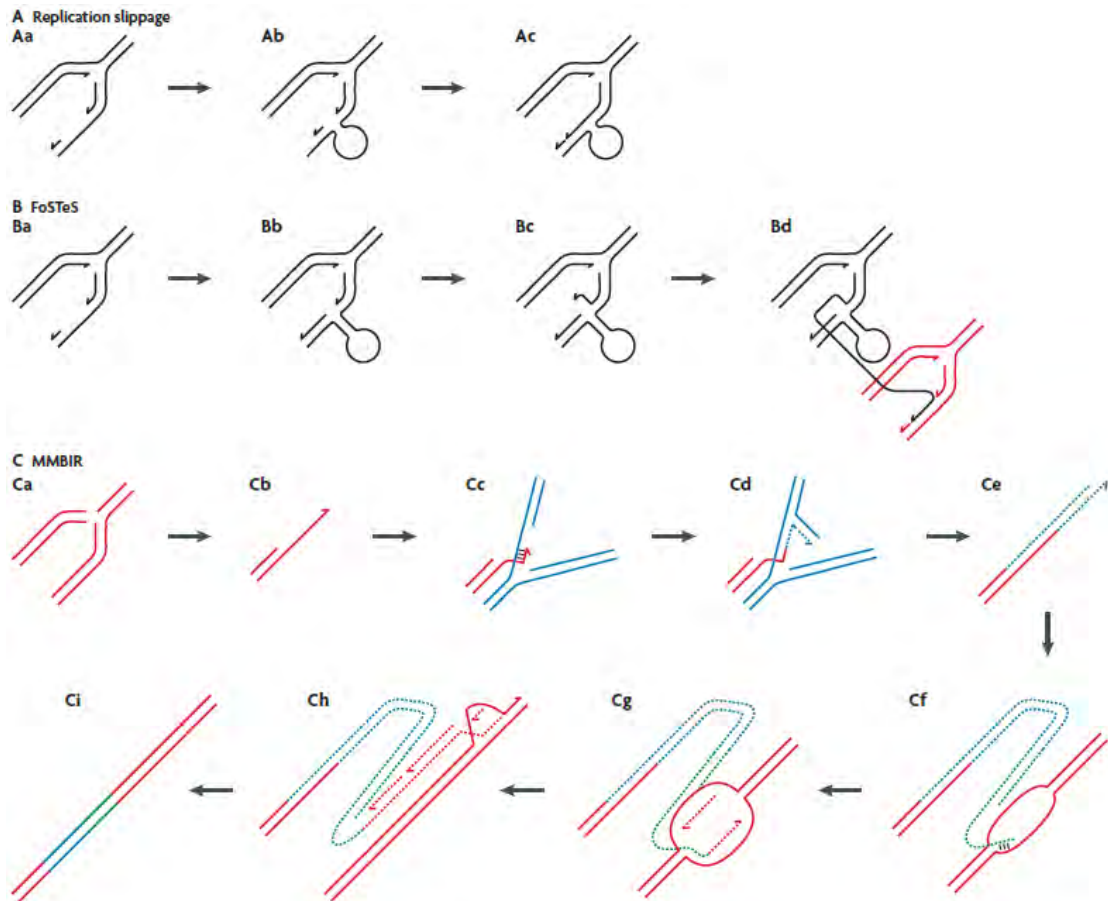
**Ολίσθηση κατά την αντιγραφή (Replication slippage) / Εναλλαγή εκμαγείου (Template switching):** Η παρουσία ταυτόσιμων αλληλουχιών μικρού μήκους στις μονόκλωνες αλυσίδες της αντιγραφικής διχάλας κατά την αντιγραφή του DNA μπορεί να οδηγήσει σε ελλείψεις και διπλασιασμούς. Κλασικό παράδειγμα τέτοιας αλληλουχίας αποτελούν τα τμήματα Okazaki. Αυτές οι δομικές μεταβολές έχουν αποδοθεί σε έναν μηχανισμό ολίσθησης της πολυμεράσης του DNA στον κλώνο που λειτουργεί ως εκμαγείο κατά τη σύνθεση του νέου κλώνου<sup>397</sup>. Σύμφωνα με μελέτες σε οργανισμούς μοντέλα, το φαινόμενο εξαρτάται σε σημαντικό βαθμό από το μέγεθος της ομολογίας και την απόσταση μεταξύ των ομόλογων περιοχών<sup>397,398</sup>. Η συχνότητα τέτοιων γεγονότων αυξάνεται από μεταλλάξεις σε γονίδια που κωδικοποιούν συστατικά απαραίτητα για την αντιγραφή, όπως η πολυμεράση του DNA, καθώς η διαταραχή της διαδικασίας έχει ως φυσικό επακόλουθο την ολίσθηση<sup>399-402</sup>. Στην περίπτωση που οι επαναλαμβανόμενες αλληλουχίες δεν ταυτίζονται εξαιτίας μονονουκλεοτιδικών αναντιστοιχιών (**mismatches**), η συχνότητα της ολίσθησης αυξάνεται παρουσία μεταλλάξεων σε συστατικά του συστήματος επιδιόρθωσής τους<sup>403</sup>. Τα δεδομένα που αφορούν ελλείψεις κρίνονται ανεπαρκή, καθώς η αφαίρεση ενός τμήματος μπορεί να επηρεάσει ένα γονίδιο ζωτικής σημασίας και να οδηγήσει το κύτταρο στο θάνατο. Όλα τα παραπάνω, σε συνδυασμό με την απουσία ενεργοποίησης μηχανισμών HR, συστήνουν έναν εναλλακτικό μηχανισμό δομικής παραλλαγής. Ωστόσο, εξαιτίας περιορισμών απόσταση των ομόλογων αλληλουχιών, ο συγκεκριμένος μηχανισμός λαμβάνει χώρα μόνο στην έκταση μιας αντιγραφικής διχάλας. Επομένως, είναι σχεδόν αδύνατο να αποτελέσει αιτία δημιουργίας CNVs στον άνθρωπο, σε αποστάσεις της τάξης των kb και Mb<sup>326</sup>.

**Αδρανοποίηση της διχάλας της αντιγραφής και εναλλαγή εκμαγείου (Fork stalling and template switching - FoSTeS):** Ο συγκεκριμένος μηχανισμός δεν περιορίζεται στο πλαίσιο μιας αντιγραφικής διχάλας, αλλά μπορεί να αλλάξει εκμαγείο μεταβαίνοντας από μια διχάλα του DNA σε μια άλλη<sup>404</sup>. Πιο συγκεκριμένα, όταν η αντιγραφική διχάλα αδρανοποιείται, το 3' άκρο της νεοσυντηθέμενης αλυσίδας μπορεί να μεταπηδήσει σε μια άλλη αντιγραφική διχάλα και να συνεχίσει εκ νέου την αντιγραφή<sup>380</sup>. Βασικό στοιχείο που τον υποστηρίζει είναι το μέσο μήκος των τμημάτων που ενισχύονταν (**amplicons**), περίπου 20kb, το οποίο είναι υπερβολικά μεγάλο στα πλαίσια της έκτασης μιας αντιγραφικής διχάλας. Ο FoSTeS παρουσιάζεται και αυτός να αξιοποιεί μικρο-ομόλογες περιοχές και γι' αυτό το λόγο κατατάσσεται ως αντιγραφικός μηχανισμός<sup>404,405</sup>. Επιπλέον, η αρχή λειτουργίας, οι ιδιότητες και η πολυπλοκότητα στην δομή των amplicons έχουν χαρακτηριστεί επαρκώς στον άνθρωπο, στο πλαίσιο σχηματισμού ελλείψεων και διπλασιασμών<sup>347,380,406,407</sup>. Ως αποτέλεσμα, ο μηχανισμός καθίσταται υπεύθυνος για την δημιουργία χρωμοσωμικών ανακατατάξεων και CNVs.



**Εικόνα 36:** Σχηματική αναπαράσταση του σχηματισμού δομικών ανακατατάξεων με τη βοήθεια μικρο-ομόλογων αλληλουχιών<sup>326,406</sup>.

**Αντιγραφικός μηχανισμός επιδιόρθωσης ρήξεων επαγόμενος από μικρο-ομολογία (Microhomology-mediated Break Induced Replication - MMBIR):** Ο μηχανισμός αυτός έχει ως χαρακτηριστικό τη θραύση ενός κλώνου κατά το σχηματισμό της αντιγραφικής δικιάλας και την δημιουργία ενός προεξέχοντος 3' άκρου<sup>104</sup>. Το άκρο αυτό στη συνέχεια μπορεί να μεταβεί σε άλλες δικιάλες αντιγραφής και να υβριδίζει σε μικρο-ομόλογες περιοχές<sup>326</sup> (**Εικόνα 36**). Τα επαναλαμβανόμενα γεγονότα της επιμήκυνσης του νεοσυντηθέμενου κλώνου και της αποδέσμευσης του από το εκμαγείο, χαρακτηριστικά για τον μηχανισμό, μπορούν να οδηγήσουν σε μια σειρά πολύπλοκων ανακατατάξεων στο γονιδίωμα. Ο MMBIR εμφανίζεται να ενεργοποιείται κυρίως σε περιπτώσεις κυτταρικού stress, όπου οι μηχανισμοί ομόλογου ανασυνδυασμού παρουσιάζονται ελαττωματικοί. Το ποσοστό της ομολογίας που απαιτείται είναι εξαιρετικά μικρό. Κατά συνέπεια, ο υβριδισμός του 3' άκρου μπορεί να προκύψει σε περιοχές ανοδικά ή καθοδικά της αρχικής στην ίδια δικιάλα, οδηγώντας σε ελλείψεις ή διπλασιασμούς με φυσιολογική ή ανάστροφη κατεύθυνση. Μικρο-ομολογίες δύνανται επίσης να υπάρχουν μεταξύ μη ομόλογων χρωμοσωμάτων, οι οποίες θα οδηγήσουν σε εκ νέου μεταθέσεις τμημάτων<sup>406</sup>. Επιπλέον, αλληλεπίδραση με το ομόλογο χρωμόσωμα αντί για την αδελφή χρωματίδα ενδέχεται να οδηγήσει σε εκτεταμένα φαινόμενα απώλειας ετεροζυγωτίας<sup>326</sup>. Ορισμένες περιπτώσεις αναφέρουν και γεγονότα απώλειας του γενετικού εντυπώματος (**imprinting**) δηλαδή του κληρονομήσιμου προτύπου μεθυλίωσης του DNA που έχει αντίκτυπο στην ρύθμιση της γονιδιακής έκφρασης<sup>384</sup>. Επικρατεί η άποψη ότι ο MMBIR ευθύνεται σε μεγάλο βαθμό για την πλειοψηφία των μη επανεμφανιζόμενων CNVs, καθώς στον σχηματισμό τους φαίνεται να συμμετέχει και ο ίδιος ο μηχανισμός αντιγραφής. Συνεπώς, παρόλο που οι δομικές ανακατατάξεις είναι λογική απόρροια των μηχανισμών που περιλαμβάνουν τη σύνθεση άκρων (**end joining mechanisms**), μπορούν να ερμηνευθούν λογικά και από μηχανισμούς όπως ο MMBIR<sup>326</sup>.



**Εικόνα 37:** Επισκόπηση των μεθόδων σχηματισμού δομικών παραλλαγών μέσω αντιγραφικών μηχανισμών επιδιόρθωσης μη ομόλογου ανασυνδυασμού<sup>326</sup>.



## Βιβλιογραφία

- 1 Sitarz, R. *et al.* Gastric cancer: epidemiology, prevention, classification, and treatment. *Cancer Manag Res* **10**, 239-248, doi:10.2147/CMAR.S149619 (2018).
- 2 Gao, J. P., Xu, W., Liu, W. T., Yan, M. & Zhu, Z. G. Tumor heterogeneity of gastric cancer: From the perspective of tumor-initiating cell. *World J Gastroenterol* **24**, 2567-2581, doi:10.3748/wjg.v24.i24.2567 (2018).
- 3 Gullo, I., Carneiro, F., Oliveira, C. & Almeida, G. M. Heterogeneity in Gastric Cancer: From Pure Morphology to Molecular Classifications. *Pathobiology* **85**, 50-63, doi:10.1159/000473881 (2018).
- 4 Hu, B. *et al.* Gastric cancer: Classification, histology and application of molecular pathology. *J Gastrointest Oncol* **3**, 251-261, doi:10.3978/j.issn.2078-6891.2012.021 (2012).
- 5 White, M. C. *et al.* Age and cancer risk: a potentially modifiable relationship. *Am J Prev Med* **46**, S7-15, doi:10.1016/j.amepre.2013.10.029 (2014).
- 6 Mahal, B. A. *et al.* Racial disparities in an aging population: the relationship between age and race in the management of African American men with high-risk prostate cancer. *J Geriatr Oncol* **5**, 352-358, doi:10.1016/j.jgo.2014.05.001 (2014).
- 7 Crew, K. D. & Neugut, A. I. Epidemiology of gastric cancer. *World J Gastroenterol* **12**, 354-362, doi:10.3748/wjg.v12.i3.354 (2006).
- 8 Ye, D. M. *et al.* Significant function and research progress of biomarkers in gastric cancer. *Oncol Lett* **19**, 17-29, doi:10.3892/ol.2019.11078 (2020).
- 9 Roukos, D. H. & Kappas, A. M. Perspectives in the treatment of gastric cancer. *Nat Clin Pract Oncol* **2**, 98-107, doi:10.1038/ncponc0099 (2005).
- 10 Ferlay, J. *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**, E359-386, doi:10.1002/ijc.29210 (2015).
- 11 Park, J. Y., Forman, D., Waskito, L. A., Yamaoka, Y. & Crabtree, J. E. Epidemiology of Helicobacter pylori and CagA-Positive Infections and Global Variations in Gastric Cancer. *Toxins (Basel)* **10**, doi:10.3390/toxins10040163 (2018).
- 12 Casamayor, M., Morlock, R., Maeda, H. & Ajani, J. Targeted literature review of the global burden of gastric cancer. *Ecancermedicalscience* **12**, 883, doi:10.3332/ecancer.2018.883 (2018).
- 13 Cheng, X. J., Lin, J. C. & Tu, S. P. Etiology and Prevention of Gastric Cancer. *Gastrointest Tumors* **3**, 25-36, doi:10.1159/000443995 (2016).
- 14 Ajani, J. A. *et al.* Gastric adenocarcinoma. *Nat Rev Dis Primers* **3**, 17036, doi:10.1038/nrdp.2017.36 (2017).
- 15 Karimi, P., Islami, F., Anandasabapathy, S., Freedman, N. D. & Kamangar, F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer Epidemiol Biomarkers Prev* **23**, 700-713, doi:10.1158/1055-9965.EPI-13-1057 (2014).
- 16 Yaghoobi, M., Bijarchi, R. & Narod, S. A. Family history and the risk of gastric cancer. *Br J Cancer* **102**, 237-242, doi:10.1038/sj.bjc.6605380 (2010).
- 17 Rivas-Ortiz, C. I., Lopez-Vidal, Y., Arredondo-Hernandez, L. J. R. & Castillo-Rojas, G. Genetic Alterations in Gastric Cancer Associated with Helicobacter pylori Infection. *Front Med (Lausanne)* **4**, 47, doi:10.3389/fmed.2017.00047 (2017).
- 18 Oliveira, C. *et al.* Genetic screening for familial gastric cancer. *Hered Cancer Clin Pract* **2**, 51-64, doi:10.1186/1897-4287-2-2-51 (2004).
- 19 Benusiglio, P. R. CDH1 germline mutations: different syndromes, same management? *Genet Med* **19**, 965-966, doi:10.1038/gim.2017.24 (2017).
- 20 Skierucha, M. *et al.* Molecular alterations in gastric cancer with special reference to the early-onset subtype. *World J Gastroenterol* **22**, 2460-2474, doi:10.3748/wjg.v22.i8.2460 (2016).

- 21 Helgason, H. *et al.* Loss-of-function variants in ATM confer risk of gastric cancer. *Nat Genet* **47**, 906-910, doi:10.1038/ng.3342 (2015).
- 22 Kula, Z. K., Zegarski, W. & Jozwicki, W. Diagnosis and treatment of early gastric cancer: experience of one center. *Prz Gastroenterol* **13**, 200-205, doi:10.5114/pg.2018.73332 (2018).
- 23 Japanese Gastric Cancer, A. Japanese classification of gastric carcinoma: 3rd English edition. *Gastric Cancer* **14**, 101-112, doi:10.1007/s10120-011-0041-5 (2011).
- 24 Van Cutsem, E., Sagaert, X., Topal, B., Haustermans, K. & Prenen, H. Gastric cancer. *Lancet* **388**, 2654-2664, doi:10.1016/S0140-6736(16)30354-3 (2016).
- 25 Venerito, M., Link, A., Rokkas, T. & Malfertheiner, P. Gastric cancer - clinical and epidemiological aspects. *Helicobacter* **21 Suppl 1**, 39-44, doi:10.1111/hel.12339 (2016).
- 26 Mukaisho, K., Nakayama, T., Hagiwara, T., Hattori, T. & Sugihara, H. Two distinct etiologies of gastric cardia adenocarcinoma: interactions among pH, *Helicobacter pylori*, and bile acids. *Front Microbiol* **6**, 412, doi:10.3389/fmicb.2015.00412 (2015).
- 27 Gong, E. J. *et al.* Characteristics of non-cardia gastric cancer with a high serum anti-*Helicobacter pylori* IgG titer and its association with diffuse-type histology. *PLoS One* **13**, e0195264, doi:10.1371/journal.pone.0195264 (2018).
- 28 Lauren, P. The Two Histological Main Types of Gastric Carcinoma: Diffuse and So-Called Intestinal-Type Carcinoma. An Attempt at a Histo-Clinical Classification. *Acta Pathol Microbiol Scand* **64**, 31-49, doi:10.1111/apm.1965.64.1.31 (1965).
- 29 Cislo, M. *et al.* Distinct molecular subtypes of gastric cancer: from Lauren to molecular pathology. *Oncotarget* **9**, 19427-19442, doi:10.18632/oncotarget.24827 (2018).
- 30 Ma, J., Shen, H., Kapesa, L. & Zeng, S. Lauren classification and individualized chemotherapy in gastric cancer. *Oncol Lett* **11**, 2959-2964, doi:10.3892/ol.2016.4337 (2016).
- 31 Henson, D. E., Dittus, C., Younes, M., Nguyen, H. & Albores-Saavedra, J. Differential trends in the intestinal and diffuse types of gastric carcinoma in the United States, 1973-2000: increase in the signet ring cell type. *Arch Pathol Lab Med* **128**, 765-770, doi:10.1043/1543-2165(2004)128<765:DTITIA>2.0.CO;2 (2004).
- 32 Qiu, M. Z. *et al.* Clinicopathological characteristics and prognostic analysis of Lauren classification in gastric adenocarcinoma in China. *J Transl Med* **11**, 58, doi:10.1186/1479-5876-11-58 (2013).
- 33 Correa, P. Gastric cancer: overview. *Gastroenterol Clin North Am* **42**, 211-217, doi:10.1016/j.gtc.2013.01.002 (2013).
- 34 Correa, P. & Piazuelo, M. B. *Helicobacter pylori* Infection and Gastric Adenocarcinoma. *US Gastroenterol Hepatol Rev* **7**, 59-64 (2011).
- 35 Correa, P., Haenszel, W., Cuello, C., Tannenbaum, S. & Archer, M. A model for gastric cancer epidemiology. *Lancet* **2**, 58-60, doi:10.1016/s0140-6736(75)90498-5 (1975).
- 36 Correa, P. Human gastric carcinogenesis: a multistep and multifactorial process--First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer Res* **52**, 6735-6740 (1992).
- 37 Correa, P. & Piazuelo, M. B. The gastric precancerous cascade. *J Dig Dis* **13**, 2-9, doi:10.1111/j.1751-2980.2011.00550.x (2012).
- 38 Malik, T. H., Sayahan, M. Y., Al Ahmed, H. A. & Hong, X. Gastric Intestinal Metaplasia: An Intermediate Precancerous Lesion in the Cascade of Gastric Carcinogenesis. *J Coll Physicians Surg Pak* **27**, 166-172 (2017).
- 39 Sandoval-Borquez, A. *et al.* Noncoding Genomics in Gastric Cancer and the Gastric Precancerous Cascade: Pathogenesis and Biomarkers. *Dis Markers* **2015**, 503762, doi:10.1155/2015/503762 (2015).

- 40 Zhang, X. Y., Zhang, P. Y. & Aboul-Soud, M. A. From inflammation to gastric cancer: Role of Helicobacter pylori. *Oncol Lett* **13**, 543-548, doi:10.3892/ol.2016.5506 (2017).
- 41 Piazzuelo, M. B., Riechelmann, R. P., Wilson, K. T. & Algood, H. M. S. Resolution of Gastric Cancer-Promoting Inflammation: A Novel Strategy for Anti-cancer Therapy. *Curr Top Microbiol Immunol* **421**, 319-359, doi:10.1007/978-3-030-15138-6\_13 (2019).
- 42 Shang, J. & Pena, A. S. Multidisciplinary approach to understand the pathogenesis of gastric cancer. *World J Gastroenterol* **11**, 4131-4139, doi:10.3748/wjg.v11.i27.4131 (2005).
- 43 Rugge, M. *et al.* Gastric dysplasia: the Padova international classification. *Am J Surg Pathol* **24**, 167-176, doi:10.1097/00000478-200002000-00001 (2000).
- 44 Riquelme, I. *et al.* Molecular classification of gastric cancer: Towards a pathway-driven targeted therapy. *Oncotarget* **6**, 24750-24779, doi:10.18632/oncotarget.4990 (2015).
- 45 Busuttill, R. A. *et al.* Role of p53 in the progression of gastric cancer. *Oncotarget* **5**, 12016-12026, doi:10.18632/oncotarget.2434 (2014).
- 46 Smith, M. G., Hold, G. L., Tahara, E. & El-Omar, E. M. Cellular and molecular aspects of gastric cancer. *World J Gastroenterol* **12**, 2979-2990, doi:10.3748/wjg.v12.i19.2979 (2006).
- 47 Naumann, M. & Crabtree, J. E. Helicobacter pylori-induced epithelial cell signalling in gastric carcinogenesis. *Trends Microbiol* **12**, 29-36, doi:10.1016/j.tim.2003.11.005 (2004).
- 48 Li, N. *et al.* Helicobacter pylori CagA Protein Negatively Regulates Autophagy and Promotes Inflammatory Response via c-Met-PI3K/Akt-mTOR Signaling Pathway. *Front Cell Infect Microbiol* **7**, 417, doi:10.3389/fcimb.2017.00417 (2017).
- 49 Churin, Y. *et al.* Helicobacter pylori CagA protein targets the c-Met receptor and enhances the motogenic response. *J Cell Biol* **161**, 249-255, doi:10.1083/jcb.200208039 (2003).
- 50 Hewitt, L. C. *et al.* KRAS status is related to histological phenotype in gastric cancer: results from a large multicentre study. *Gastric Cancer* **22**, 1193-1203, doi:10.1007/s10120-019-00972-6 (2019).
- 51 Yasui, W. *et al.* Molecular diagnosis of gastric cancer: present and future. *Gastric Cancer* **4**, 113-121, doi:10.1007/pl00011733 (2001).
- 52 Tamura, G. Alterations of tumor suppressor and tumor-related genes in the development and progression of gastric cancer. *World J Gastroenterol* **12**, 192-198, doi:10.3748/wjg.v12.i2.192 (2006).
- 53 Yasui, W., Sentani, K., Motoshita, J. & Nakayama, H. Molecular pathobiology of gastric cancer. *Scand J Surg* **95**, 225-231, doi:10.1177/145749690609500403 (2006).
- 54 Ohtsuka, J. *et al.* Functional loss of p53 cooperates with the in vivo microenvironment to promote malignant progression of gastric cancers. *Sci Rep* **8**, 2291, doi:10.1038/s41598-018-20572-1 (2018).
- 55 Morgan, C. *et al.* Detection of p53 mutations in precancerous gastric tissue. *Br J Cancer* **89**, 1314-1319, doi:10.1038/sj.bjc.6601302 (2003).
- 56 Li, J. H., Shi, X. Z., Lv, S., Liu, M. & Xu, G. W. Effect of Helicobacter pylori infection on p53 expression of gastric mucosa and adenocarcinoma with microsatellite instability. *World J Gastroenterol* **11**, 4363-4366, doi:10.3748/wjg.v11.i28.4363 (2005).
- 57 Fang, Z. *et al.* APC gene deletions in gastric adenocarcinomas in a Chinese population: a correlation with tumour progression. *Clin Transl Oncol* **14**, 60-65, doi:10.1007/s12094-012-0762-x (2012).
- 58 Nakatsuru, S. *et al.* Somatic mutations of the APC gene in precancerous lesion of the stomach. *Hum Mol Genet* **2**, 1463-1465, doi:10.1093/hmg/2.9.1463 (1993).

- 59 Agnes, A., Estrella, J. S. & Badgwell, B. The significance of a nineteenth century definition in the era of genomics: linitis plastica. *World J Surg Oncol* **15**, 123, doi:10.1186/s12957-017-1187-3 (2017).
- 60 Barber, M. *et al.* Mechanisms and sequelae of E-cadherin silencing in hereditary diffuse gastric cancer. *J Pathol* **216**, 295-306, doi:10.1002/path.2426 (2008).
- 61 Luo, W., Fedda, F., Lynch, P. & Tan, D. CDH1 Gene and Hereditary Diffuse Gastric Cancer Syndrome: Molecular and Histological Alterations and Implications for Diagnosis And Treatment. *Front Pharmacol* **9**, 1421, doi:10.3389/fphar.2018.01421 (2018).
- 62 Liu, X. & Chu, K. M. E-cadherin and gastric cancer: cause, consequence, and applications. *Biomed Res Int* **2014**, 637308, doi:10.1155/2014/637308 (2014).
- 63 Moreira, A. M. *et al.* The Extracellular Matrix: An Accomplice in Gastric Cancer Development and Progression. *Cells* **9**, doi:10.3390/cells9020394 (2020).
- 64 Fitzgerald, R. C. *et al.* Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. *J Med Genet* **47**, 436-444, doi:10.1136/jmg.2009.074237 (2010).
- 65 Charlton, A. *et al.* Hereditary diffuse gastric cancer: predominance of multiple foci of signet ring cell carcinoma in distal stomach and transitional zone. *Gut* **53**, 814-820, doi:10.1136/gut.2002.010447 (2004).
- 66 Grady, W. M. *et al.* Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. *Nat Genet* **26**, 16-17, doi:10.1038/79120 (2000).
- 67 Humar, B. *et al.* Association of CDH1 haplotypes with susceptibility to sporadic diffuse gastric cancer. *Oncogene* **21**, 8192-8195, doi:10.1038/sj.onc.1205921 (2002).
- 68 Wang, Q., Liu, G. & Hu, C. Molecular Classification of Gastric Adenocarcinoma. *Gastroenterology Res* **12**, 275-282, doi:10.14740/gr1187 (2019).
- 69 Cancer Genome Atlas Research, N. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202-209, doi:10.1038/nature13480 (2014).
- 70 Tang, W. *et al.* Epstein-barr virus infected gastric adenocarcinoma expresses latent and lytic viral transcripts and has a distinct human gene expression profile. *Infect Agent Cancer* **7**, 21, doi:10.1186/1750-9378-7-21 (2012).
- 71 Camargo, M. C. *et al.* Improved survival of gastric cancer with tumour Epstein-Barr virus positivity: an international pooled analysis. *Gut* **63**, 236-243, doi:10.1136/gutjnl-2013-304531 (2014).
- 72 Arcaro, A. & Guerreiro, A. S. The phosphoinositide 3-kinase pathway in human cancer: genetic alterations and therapeutic implications. *Curr Genomics* **8**, 271-306, doi:10.2174/138920207782446160 (2007).
- 73 Hoadley, K. A. *et al.* Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell* **158**, 929-944, doi:10.1016/j.cell.2014.06.049 (2014).
- 74 Kandath, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333-339, doi:10.1038/nature12634 (2013).
- 75 Wu, H. *et al.* MiR-135a targets JAK2 and inhibits gastric cancer cell proliferation. *Cancer Biol Ther* **13**, 281-288, doi:10.4161/cbt.18943 (2012).
- 76 Thomas, S. J., Snowden, J. A., Zeidler, M. P. & Danson, S. J. The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours. *Br J Cancer* **113**, 365-371, doi:10.1038/bjc.2015.233 (2015).
- 77 McDermott, D. F. & Atkins, M. B. PD-1 as a potential target in cancer therapy. *Cancer Med* **2**, 662-673, doi:10.1002/cam4.106 (2013).
- 78 Sun, J. *et al.* PD-L1 expression analysis in gastric carcinoma tissue and blocking of tumor-associated PD-L1 signaling by two functional monoclonal antibodies. *Tissue Antigens* **69**, 19-27, doi:10.1111/j.1399-0039.2006.00701.x (2007).

- 79 Kim, J. Y. *et al.* Microsatellite instability status in gastric cancer: a reappraisal of its clinical significance and relationship with mucin phenotypes. *Korean J Pathol* **47**, 28-35, doi:10.4132/KoreanJPathol.2013.47.1.28 (2013).
- 80 Falchetti, M. *et al.* Gastric cancer with high-level microsatellite instability: target gene mutations, clinicopathologic features, and long-term survival. *Hum Pathol* **39**, 925-932, doi:10.1016/j.humpath.2007.10.024 (2008).
- 81 Choi, Y. Y. *et al.* Is microsatellite instability a prognostic marker in gastric cancer? A systematic review with meta-analysis. *J Surg Oncol* **110**, 129-135, doi:10.1002/jso.23618 (2014).
- 82 Haron, N. H. *et al.* Microsatellite Instability and Altered Expressions of MLH1 and MSH2 in Gastric Cancer. *Asian Pac J Cancer Prev* **20**, 509-517, doi:10.31557/APJCP.2019.20.2.509 (2019).
- 83 Yamamoto, H., Sawai, H. & Perucho, M. Frameshift somatic mutations in gastrointestinal cancer of the microsatellite mutator phenotype. *Cancer Res* **57**, 4420-4426 (1997).
- 84 Gazvoda, B. *et al.* Genetic changes in Slovenian patients with gastric adenocarcinoma evaluated in terms of microsatellite DNA. *Eur J Gastroenterol Hepatol* **19**, 1082-1089, doi:10.1097/MEG.0b013e3282f13cf9 (2007).
- 85 McLean, M. H. & El-Omar, E. M. Genetics of gastric cancer. *Nat Rev Gastroenterol Hepatol* **11**, 664-674, doi:10.1038/nrgastro.2014.143 (2014).
- 86 French, A. J. *et al.* Allelic imbalance of 8p indicates poor survival in gastric cancer. *J Mol Diagn* **6**, 243-252, doi:10.1016/S1525-1578(10)60517-X (2004).
- 87 Wang, K. *et al.* Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat Genet* **46**, 573-582, doi:10.1038/ng.2983 (2014).
- 88 Kakiuchi, M. *et al.* Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nat Genet* **46**, 583-587, doi:10.1038/ng.2984 (2014).
- 89 Murray, D., Horgan, G., Macmathuna, P. & Doran, P. NET1-mediated RhoA activation facilitates lysophosphatidic acid-induced cell migration and invasion in gastric cancer. *Br J Cancer* **99**, 1322-1329, doi:10.1038/sj.bjc.6604688 (2008).
- 90 Liu, N. *et al.* Reversal of the malignant phenotype of gastric cancer cells by inhibition of RhoA expression and activity. *Clin Cancer Res* **10**, 6239-6247, doi:10.1158/1078-0432.CCR-04-0242 (2004).
- 91 Lin, M. T. *et al.* IL-6 induces AGS gastric cancer cell invasion via activation of the c-Src/RhoA/ROCK signaling pathway. *Int J Cancer* **120**, 2600-2608, doi:10.1002/ijc.22599 (2007).
- 92 Yao, F. *et al.* Recurrent Fusion Genes in Gastric Cancer: CLDN18-ARHGAP26 Induces Loss of Epithelial Integrity. *Cell reports* **12**, 272-285, doi:10.1016/j.celrep.2015.06.020 (2015).
- 93 Ben-David, U. & Amon, A. Context is everything: aneuploidy in cancer. *Nat Rev Genet* **21**, 44-62, doi:10.1038/s41576-019-0171-x (2020).
- 94 Ayed-Guerfali, D. B. *et al.* Expression of APC, beta-catenin and E-cadherin in Tunisian patients with gastric adenocarcinoma: clinical significance. *Tumour Biol* **35**, 1775-1783, doi:10.1007/s13277-013-1236-7 (2014).
- 95 Liang, L., Fang, J. Y. & Xu, J. Gastric cancer and gene copy number variation: emerging cancer drivers for targeted therapy. *Oncogene* **35**, 1475-1482, doi:10.1038/onc.2015.209 (2016).
- 96 Morishita, A., Gong, J. & Masaki, T. Targeting receptor tyrosine kinases in gastric cancer. *World J Gastroenterol* **20**, 4536-4545, doi:10.3748/wjg.v20.i16.4536 (2014).

- 97 Ooi, A. *et al.* Gene amplification of CCNE1, CCND1, and CDK6 in gastric cancers detected by multiplex ligation-dependent probe amplification and fluorescence in situ hybridization. *Hum Pathol* **61**, 58-67, doi:10.1016/j.humpath.2016.10.025 (2017).
- 98 Sun, Z. & Yan, B. Multiple roles and regulatory mechanisms of the transcription factor GATA6 in human cancers. *Clin Genet* **97**, 64-72, doi:10.1111/cge.13630 (2020).
- 99 Garraway, L. A. & Sellers, W. R. Lineage dependency and lineage-survival oncogenes in human cancer. *Nat Rev Cancer* **6**, 593-602, doi:10.1038/nrc1947 (2006).
- 100 Chia, N. Y. *et al.* Regulatory crosstalk between lineage-survival oncogenes KLF5, GATA4 and GATA6 cooperatively promotes gastric cancer development. *Gut* **64**, 707-719, doi:10.1136/gutjnl-2013-306596 (2015).
- 101 Deng, N. *et al.* A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. *Gut* **61**, 673-684, doi:10.1136/gutjnl-2011-301839 (2012).
- 102 Alessandrini, L., Manchi, M., Italia, F., Perin, T. & Canzonieri, V. in *Gastric Cancer In The Precision Medicine Era: Diagnosis and Therapy* (eds Vincenzo Canzonieri & Antonio Giordano) 155-172 (Springer International Publishing, 2019).
- 103 Zhang, F. & Lupski, J. R. Non-coding genetic variants in human disease. *Hum Mol Genet* **24**, R102-110, doi:10.1093/hmg/ddv259 (2015).
- 104 Carvalho, C. M. & Lupski, J. R. Mechanisms underlying structural variant formation in genomic disorders. *Nat Rev Genet* **17**, 224-238, doi:10.1038/nrg.2015.25 (2016).
- 105 Alkan, C., Coe, B. P. & Eichler, E. E. Genome structural variation discovery and genotyping. *Nat Rev Genet* **12**, 363-376, doi:10.1038/nrg2958 (2011).
- 106 Escaramis, G., Docampo, E. & Rabionet, R. A decade of structural variants: description, history and methods to detect structural variation. *Brief Funct Genomics* **14**, 305-314, doi:10.1093/bfgp/elv014 (2015).
- 107 Zhang, F., Gu, W., Hurles, M. E. & Lupski, J. R. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* **10**, 451-481, doi:10.1146/annurev.genom.9.081307.164217 (2009).
- 108 Glover, T. W., Wilson, T. E. & Arlt, M. F. Fragile sites in cancer: more than meets the eye. *Nat Rev Cancer* **17**, 489-501, doi:10.1038/nrc.2017.52 (2017).
- 109 Guilherme, R. S. *et al.* Mechanisms of ring chromosome formation, ring instability and clinical consequences. *BMC Med Genet* **12**, 171, doi:10.1186/1471-2350-12-171 (2011).
- 110 Rodriguez-Santiago, B. *et al.* Mosaic uniparental disomies and aneuploidies as large structural variants of the human genome. *Am J Hum Genet* **87**, 129-138, doi:10.1016/j.ajhg.2010.06.002 (2010).
- 111 Zarrei, M., MacDonald, J. R., Merico, D. & Scherer, S. W. A copy number variation map of the human genome. *Nat Rev Genet* **16**, 172-183, doi:10.1038/nrg3871 (2015).
- 112 Piotrowski, A. *et al.* Somatic mosaicism for copy number variation in differentiated human tissues. *Hum Mutat* **29**, 1118-1124, doi:10.1002/humu.20815 (2008).
- 113 Beckmann, J. S., Estivill, X. & Antonarakis, S. E. Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet* **8**, 639-646, doi:10.1038/nrg2149 (2007).
- 114 Sudmant, P. H. *et al.* Global diversity, population stratification, and selection of human copy-number variation. *Science* **349**, aab3761, doi:10.1126/science.aab3761 (2015).
- 115 Venkatapoorna, C. M. K. *et al.* Association of Salivary Amylase (AMY1) Gene Copy Number with Obesity in Alabama Elementary School Children. *Nutrients* **11**, doi:10.3390/nu11061379 (2019).
- 116 Perry, G. H. *et al.* Diet and the evolution of human amylase gene copy number variation. *Nat Genet* **39**, 1256-1260, doi:10.1038/ng2123 (2007).

- 117 Tian, D. *et al.* Single-nucleotide mutation rate increases close to insertions/deletions  
in eukaryotes. *Nature* **455**, 105-108, doi:10.1038/nature07175 (2008).
- 118 de Smith, A. J., Walters, R. G., Froguel, P. & Blakemore, A. I. Human genes involved in  
copy number variation: mechanisms of origin, functional effects and implications for  
disease. *Cytogenet Genome Res* **123**, 17-26, doi:10.1159/000184688 (2008).
- 119 Myers, S., Freeman, C., Auton, A., Donnelly, P. & McVean, G. A common sequence  
motif associated with recombination hot spots and genome instability in humans. *Nat*  
*Genet* **40**, 1124-1129, doi:10.1038/ng.213 (2008).
- 120 Fu, W., Zhang, F., Wang, Y., Gu, X. & Jin, L. Identification of copy number variation  
hotspots in human populations. *Am J Hum Genet* **87**, 494-504,  
doi:10.1016/j.ajhg.2010.09.006 (2010).
- 121 Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**,  
444-454, doi:10.1038/nature05329 (2006).
- 122 Cooper, G. M., Nickerson, D. A. & Eichler, E. E. Mutational and selective effects on  
copy-number variants in the human genome. *Nat Genet* **39**, S22-29,  
doi:10.1038/ng2054 (2007).
- 123 Nguyen, D. Q., Webber, C. & Ponting, C. P. Bias of selection on human copy-number  
variants. *PLoS Genet* **2**, e20, doi:10.1371/journal.pgen.0020020 (2006).
- 124 Grimwood, J. *et al.* The DNA sequence and biology of human chromosome 19. *Nature*  
**428**, 529-535, doi:10.1038/nature02399 (2004).
- 125 Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Philos*  
*Trans R Soc Lond B Biol Sci* **355**, 1563-1572, doi:10.1098/rstb.2000.0717 (2000).
- 126 Korbelt, J. O. *et al.* Paired-end mapping reveals extensive structural variation in the  
human genome. *Science* **318**, 420-426, doi:10.1126/science.1149504 (2007).
- 127 Feuk, L., Carson, A. R. & Scherer, S. W. Structural variation in the human genome. *Nat*  
*Rev Genet* **7**, 85-97, doi:10.1038/nrg1767 (2006).
- 128 Wong, K. K. *et al.* A comprehensive analysis of common copy-number variations in the  
human genome. *Am J Hum Genet* **80**, 91-104, doi:10.1086/510560 (2007).
- 129 de Smith, A. J. *et al.* Array CGH analysis of copy number variation identifies 1284 new  
genes variant in healthy white males: implications for association studies of complex  
diseases. *Hum Mol Genet* **16**, 2783-2794, doi:10.1093/hmg/ddm208 (2007).
- 130 Conrad, D. F., Andrews, T. D., Carter, N. P., Hurler, M. E. & Pritchard, J. K. A high-  
resolution survey of deletion polymorphism in the human genome. *Nat Genet* **38**, 75-  
81, doi:10.1038/ng1697 (2006).
- 131 Rice, A. M. & McLysaght, A. Dosage sensitivity is a major determinant of human copy  
number variant pathogenicity. *Nat Commun* **8**, 14366, doi:10.1038/ncomms14366  
(2017).
- 132 Rice, A. M. & McLysaght, A. Dosage-sensitive genes in evolution and disease. *BMC Biol*  
**15**, 78, doi:10.1186/s12915-017-0418-y (2017).
- 133 Stranger, B. E. *et al.* Relative impact of nucleotide and copy number variation on gene  
expression phenotypes. *Science* **315**, 848-853, doi:10.1126/science.1136678 (2007).
- 134 McCarroll, S. A. *et al.* Common deletion polymorphisms in the human genome. *Nat*  
*Genet* **38**, 86-92, doi:10.1038/ng1696 (2006).
- 135 Miller, D. W. *et al.* Alpha-synuclein in blood and brain from familial Parkinson disease  
with SNCA locus triplication. *Neurology* **62**, 1835-1838,  
doi:10.1212/01.wnl.0000127517.33208.f4 (2004).
- 136 Al-Balool, H. H. *et al.* Post-transcriptional exon shuffling events in humans can be  
evolutionarily conserved and abundant. *Genome Res* **21**, 1788-1799,  
doi:10.1101/gr.116442.110 (2011).

- 137 Park, E., Pan, Z., Zhang, Z., Lin, L. & Xing, Y. The Expanding Landscape of Alternative Splicing Variation in Human Populations. *Am J Hum Genet* **102**, 11-26, doi:10.1016/j.ajhg.2017.11.002 (2018).
- 138 Chen, X. & Zhang, J. The Genomic Landscape of Position Effects on Protein Expression Level and Noise in Yeast. *Cell Syst* **2**, 347-354, doi:10.1016/j.cels.2016.03.009 (2016).
- 139 Rodriguez-Revenga, L., Mila, M., Rosenberg, C., Lamb, A. & Lee, C. Structural variation in the human genome: the impact of copy number variants on clinical diagnosis. *Genet Med* **9**, 600-606, doi:10.1097/gim.0b013e318149e1e3 (2007).
- 140 Kelsey, K. J. P. & Clark, A. G. Variation in Position Effect Variegation Within a Natural Population. *Genetics* **207**, 1157-1166, doi:10.1534/genetics.117.300306 (2017).
- 141 Gamazon, E. R. & Stranger, B. E. The impact of human copy number variation on gene expression. *Brief Funct Genomics* **14**, 352-357, doi:10.1093/bfgp/elv017 (2015).
- 142 Nowakowska, B. Clinical interpretation of copy number variants in the human genome. *J Appl Genet* **58**, 449-457, doi:10.1007/s13353-017-0407-4 (2017).
- 143 Westerfield, L., Darilek, S. & van den Veyver, I. B. Counseling Challenges with Variants of Uncertain Significance and Incidental Findings in Prenatal Genetic Screening and Diagnosis. *J Clin Med* **3**, 1018-1032, doi:10.3390/jcm3031018 (2014).
- 144 Vittori, A. *et al.* Copy-number variation of the neuronal glucose transporter gene SLC2A3 and age of onset in Huntington's disease. *Hum Mol Genet* **23**, 3129-3137, doi:10.1093/hmg/ddu022 (2014).
- 145 Rovelet-Lecrux, A. *et al.* APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet* **38**, 24-26, doi:10.1038/ng1718 (2006).
- 146 Singleton, A. B. *et al.* alpha-Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841, doi:10.1126/science.1090278 (2003).
- 147 Mishra, S. & Whetstine, J. R. Different Facets of Copy Number Changes: Permanent, Transient, and Adaptive. *Mol Cell Biol* **36**, 1050-1063, doi:10.1128/MCB.00652-15 (2016).
- 148 Beroukhi, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899-905, doi:10.1038/nature08822 (2010).
- 149 International Cancer Genome, C. *et al.* International network of cancer genome projects. *Nature* **464**, 993-998, doi:10.1038/nature08987 (2010).
- 150 Shao, X. *et al.* Copy number variation is highly correlated with differential gene expression: a pan-cancer study. *BMC Med Genet* **20**, 175, doi:10.1186/s12881-019-0909-5 (2019).
- 151 Lopez-Gines, C. *et al.* New pattern of EGFR amplification in glioblastoma and the relationship of gene copy number with gene expression profile. *Mod Pathol* **23**, 856-865, doi:10.1038/modpathol.2010.62 (2010).
- 152 Yoda, H. *et al.* Direct Targeting of MYCN Gene Amplification by Site-Specific DNA Alkylation in Neuroblastoma. *Cancer Res* **79**, 830-840, doi:10.1158/0008-5472.CAN-18-1198 (2019).
- 153 Lee, S. *et al.* Acute myeloid leukemia with MYC amplification in the homogeneous staining regions and double minutes. *Cancer Genet Cytogenet* **192**, 96-98, doi:10.1016/j.cancergencyto.2009.04.003 (2009).
- 154 Lee, J. *et al.* Detection of ERBB2 (HER2) Gene Amplification Events in Cell-Free DNA and Response to Anti-HER2 Agents in a Large Asian Cancer Patient Cohort. *Front Oncol* **9**, 212, doi:10.3389/fonc.2019.00212 (2019).
- 155 Sendor, A. B. *et al.* Von Hippel-Lindau status influences phenotype of liver cancers arising from PTEN loss. *Gastrointest Cancer* **5**, 61-71, doi:10.2147/GICTT.S72274 (2015).



- 156 Hollander, M. C., Blumenthal, G. M. & Dennis, P. A. PTEN loss in the continuum of common cancers, rare syndromes and mouse models. *Nat Rev Cancer* **11**, 289-301, doi:10.1038/nrc3037 (2011).
- 157 Pagliarini, R., Shao, W. & Sellers, W. R. Oncogene addiction: pathways of therapeutic response, resistance, and road maps toward a cure. *EMBO Rep* **16**, 280-296, doi:10.15252/embr.201439949 (2015).
- 158 Torti, D. & Trusolino, L. Oncogene addiction as a foundational rationale for targeted anti-cancer therapy: promises and perils. *EMBO Mol Med* **3**, 623-636, doi:10.1002/emmm.201100176 (2011).
- 159 Weinstein, I. B. & Joe, A. K. Mechanisms of disease: Oncogene addiction--a rationale for molecular targeting in cancer therapy. *Nat Clin Pract Oncol* **3**, 448-457, doi:10.1038/nconc0558 (2006).
- 160 Vivanco, I. Targeting molecular addictions in cancer. *Br J Cancer* **111**, 2033-2038, doi:10.1038/bjc.2014.461 (2014).
- 161 Krijgsman, O., Carvalho, B., Meijer, G. A., Steenbergen, R. D. & Ylstra, B. Focal chromosomal copy number aberrations in cancer--Needles in a genome haystack. *Biochim Biophys Acta* **1843**, 2698-2704, doi:10.1016/j.bbamcr.2014.08.001 (2014).
- 162 Zhang, X. *et al.* Identification of focally amplified lineage-specific super-enhancers in human epithelial cancers. *Nat Genet* **48**, 176-182, doi:10.1038/ng.3470 (2016).
- 163 Williams, M. M. *et al.* Therapeutic inhibition of Mcl-1 blocks cell survival in estrogen receptor-positive breast cancers. *Oncotarget* **10**, 5389-5402, doi:10.18632/oncotarget.27070 (2019).
- 164 Schmidt, T. M. *et al.* Gain of Chromosome 1q is associated with early progression in multiple myeloma patients treated with lenalidomide, bortezomib, and dexamethasone. *Blood Cancer J* **9**, 94, doi:10.1038/s41408-019-0254-0 (2019).
- 165 Shaughnessy, J. Amplification and overexpression of CKS1B at chromosome band 1q21 is associated with reduced levels of p27Kip1 and an aggressive clinical course in multiple myeloma. *Hematology* **10 Suppl 1**, 117-126, doi:10.1080/10245330512331390140 (2005).
- 166 Fonseca, R. *et al.* Prognostic value of chromosome 1q21 gain by fluorescent in situ hybridization and increase CKS1B expression in myeloma. *Leukemia* **20**, 2034-2040, doi:10.1038/sj.leu.2404403 (2006).
- 167 Yin, L., Kosugi, M. & Kufe, D. Inhibition of the MUC1-C oncoprotein induces multiple myeloma cell death by down-regulating TIGAR expression and depleting NADPH. *Blood* **119**, 810-816, doi:10.1182/blood-2011-07-369686 (2012).
- 168 Fan, F. *et al.* Targeting Mcl-1 for multiple myeloma (MM) therapy: drug-induced generation of Mcl-1 fragment Mcl-1(128-350) triggers MM cell death via c-Jun upregulation. *Cancer Lett* **343**, 286-294, doi:10.1016/j.canlet.2013.09.042 (2014).
- 169 Patch, A. M. *et al.* Whole-genome characterization of chemoresistant ovarian cancer. *Nature* **521**, 489-494, doi:10.1038/nature14410 (2015).
- 170 Takano, M. *et al.* Analyses by comparative genomic hybridization of genes relating with cisplatin-resistance in ovarian cancer. *Hum Cell* **14**, 267-271 (2001).
- 171 Etemadmoghadam, D. *et al.* Resistance to CDK2 inhibitors is associated with selection of polyploid cells in CCNE1-amplified ovarian cancer. *Clin Cancer Res* **19**, 5960-5971, doi:10.1158/1078-0432.CCR-13-1337 (2013).
- 172 Yabuki, N. *et al.* Gene amplification and expression in lung cancer cells with acquired paclitaxel resistance. *Cancer Genet Cytogenet* **173**, 1-9, doi:10.1016/j.cancergencyto.2006.07.020 (2007).
- 173 Yoda, Y. *et al.* Integrated analysis of cancer-related pathways affected by genetic and epigenetic alterations in gastric cancer. *Gastric Cancer* **18**, 65-76, doi:10.1007/s10120-014-0348-0 (2015).

- 174 Takeshima, H. *et al.* Frequent involvement of chromatin remodeler alterations in gastric field cancerization. *Cancer Lett* **357**, 328-338, doi:10.1016/j.canlet.2014.11.038 (2015).
- 175 Katoh, M. Therapeutics targeting angiogenesis: genetics and epigenetics, extracellular miRNAs and signaling networks (Review). *Int J Mol Med* **32**, 763-767, doi:10.3892/ijmm.2013.1444 (2013).
- 176 Hudler, P. Genetic aspects of gastric cancer instability. *ScientificWorldJournal* **2012**, 761909, doi:10.1100/2012/761909 (2012).
- 177 Kiyose, S. *et al.* Chromogenic in situ hybridization (CISH) to detect HER2 gene amplification in breast and gastric cancer: comparison with immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH). *Pathol Int* **62**, 728-734, doi:10.1111/j.1440-1827.2012.02862.x (2012).
- 178 Krepischki, A. C., Pearson, P. L. & Rosenberg, C. Germline copy number variations and cancer predisposition. *Future Oncol* **8**, 441-450, doi:10.2217/fon.12.34 (2012).
- 179 Leary, R. J. *et al.* Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc Natl Acad Sci U S A* **105**, 16224-16229, doi:10.1073/pnas.0808041105 (2008).
- 180 Samuels, Y. *et al.* High frequency of mutations of the PIK3CA gene in human cancers. *Science* **304**, 554, doi:10.1126/science.1096502 (2004).
- 181 Shi, J. *et al.* Highly frequent PIK3CA amplification is associated with poor prognosis in gastric cancer. *BMC Cancer* **12**, 50, doi:10.1186/1471-2407-12-50 (2012).
- 182 Yang, Q. *et al.* Concomitant PIK3CA amplification and RASSF1A or PAX6 hypermethylation predict worse survival in gastric cancer. *Clin Biochem* **47**, 111-116 (2014).
- 183 Yoshida, S. *et al.* Gene amplification of ribosomal protein S6 kinase-1 and -2 in gastric cancer. *Anticancer Res* **33**, 469-475 (2013).
- 184 Shinmura, K. *et al.* TNK2 gene amplification is a novel predictor of a poor prognosis in patients with gastric cancer. *J Surg Oncol* **109**, 189-197, doi:10.1002/jso.23482 (2014).
- 185 Li, Z., Li, W., Song, L. & Zhu, W. Cilia, adenomatous polyposis coli and associated diseases. *Oncogene* **31**, 1475-1483, doi:10.1038/onc.2011.351 (2012).
- 186 Tsai, P. C. *et al.* The association between DNA copy number aberrations at chromosome 5q22 and gastric cancer. *PLoS One* **9**, e106624, doi:10.1371/journal.pone.0106624 (2014).
- 187 Liu, X., Newton, R. C. & Scherle, P. A. Developing c-MET pathway inhibitors for cancer therapy: progress and challenges. *Trends Mol Med* **16**, 37-45, doi:10.1016/j.molmed.2009.11.005 (2010).
- 188 Jung, K. H., Park, B. H. & Hong, S. S. Progress in cancer therapy targeting c-Met signaling pathway. *Arch Pharm Res* **35**, 595-604, doi:10.1007/s12272-012-0402-6 (2012).
- 189 Gherardi, E., Birchmeier, W., Birchmeier, C. & Vande Woude, G. Targeting MET in cancer: rationale and progress. *Nat Rev Cancer* **12**, 89-103, doi:10.1038/nrc3205 (2012).
- 190 Yarden, Y. & Pines, G. The ERBB network: at last, cancer therapy meets systems biology. *Nat Rev Cancer* **12**, 553-563, doi:10.1038/nrc3309 (2012).
- 191 Hong, L., Han, Y. & Brain, L. The role of epidermal growth factor receptor in prognosis and treatment of gastric cancer. *Expert Rev Gastroenterol Hepatol* **8**, 111-117, doi:10.1586/17474124.2014.844648 (2014).
- 192 Higaki, E. *et al.* Gene copy number gain of EGFR is a poor prognostic biomarker in gastric cancer: evaluation of 855 patients with bright-field dual in situ hybridization (DISH) method. *Gastric Cancer* **19**, 63-73, doi:10.1007/s10120-014-0449-9 (2016).

- 193 Zhang, L. *et al.* The impact of C-MYC gene expression on gastric cancer cell. *Mol Cell Biochem* **344**, 125-135, doi:10.1007/s11010-010-0536-0 (2010).
- 194 Herold, S., Herkert, B. & Eilers, M. Facilitating replication under stress: an oncogenic function of MYC? *Nat Rev Cancer* **9**, 441-444, doi:10.1038/nrc2640 (2009).
- 195 Zhao, S. *et al.* Expression of OCT4 pseudogenes in human tumours: lessons from glioma and breast carcinoma. *J Pathol* **223**, 672-682, doi:10.1002/path.2827 (2011).
- 196 Hayashi, H. *et al.* The OCT4 pseudogene POU5F1B is amplified and promotes an aggressive phenotype in gastric cancer. *Oncogene* **34**, 199-208, doi:10.1038/onc.2013.547 (2015).
- 197 Shim, J. H. *et al.* The effect of Helicobacter pylori CagA on the HER-2 copy number and expression in gastric cancer. *Gene* **546**, 288-296, doi:10.1016/j.gene.2014.05.064 (2014).
- 198 Sheng, W. Q. *et al.* HER2 status in gastric cancers: a retrospective analysis from four Chinese representative clinical centers and assessment of its prognostic significance. *Ann Oncol* **24**, 2360-2364, doi:10.1093/annonc/mdt232 (2013).
- 199 Matsuoka, T. & Yashiro, M. Recent advances in the HER2 targeted therapy of gastric cancer. *World J Clin Cases* **3**, 42-51, doi:10.12998/wjcc.v3.i1.42 (2015).
- 200 Xu, J. *et al.* Unequal prognostic potentials of p53 gain-of-function mutations in human cancers associate with drug-metabolizing activity. *Cell Death Dis* **5**, e1108, doi:10.1038/cddis.2014.75 (2014).
- 201 Ozaki, T. & Nakagawara, A. p53: the attractive tumor suppressor in the cancer research field. *J Biomed Biotechnol* **2011**, 603925, doi:10.1155/2011/603925 (2011).
- 202 Wang, J. *et al.* ArhGAP30 promotes p53 acetylation and function in colorectal cancer. *Nat Commun* **5**, 4735, doi:10.1038/ncomms5735 (2014).
- 203 Suzuki, K. & Matsubara, H. Recent advances in p53 research and cancer treatment. *J Biomed Biotechnol* **2011**, 978312, doi:10.1155/2011/978312 (2011).
- 204 Bellini, M. F., Cadamuro, A. C., Succì, M., Proenca, M. A. & Silva, A. E. Alterations of the TP53 gene in gastric and esophageal carcinogenesis. *J Biomed Biotechnol* **2012**, 891961, doi:10.1155/2012/891961 (2012).
- 205 Sonoda, A. *et al.* Genetic lineages of undifferentiated-type gastric carcinomas analysed by unsupervised clustering of genomic DNA microarray data. *BMC Med Genomics* **6**, 25, doi:10.1186/1755-8794-6-25 (2013).
- 206 Silva, T. C. *et al.* hTERT, MYC and TP53 deregulation in gastric preneoplastic lesions. *BMC Gastroenterol* **12**, 85, doi:10.1186/1471-230X-12-85 (2012).
- 207 Leal, M. F. *et al.* MYC, TP53, and chromosome 17 copy-number alterations in multiple gastric cancer cell lines and in their parental primary tumors. *J Biomed Biotechnol* **2011**, 631268, doi:10.1155/2011/631268 (2011).
- 208 Fassan, M. *et al.* High-throughput mutation profiling identifies novel molecular dysregulation in high-grade intraepithelial neoplasia and early gastric cancers. *Gastric Cancer* **17**, 442-449, doi:10.1007/s10120-013-0315-1 (2014).
- 209 Liao, L. M. *et al.* Mitochondrial DNA copy number and risk of gastric cancer: a report from the Shanghai Women's Health Study. *Cancer Epidemiol Biomarkers Prev* **20**, 1944-1949, doi:10.1158/1055-9965.EPI-11-0379 (2011).
- 210 Weinberg, S. E. & Chandel, N. S. Targeting mitochondria metabolism for cancer therapy. *Nat Chem Biol* **11**, 9-15, doi:10.1038/nchembio.1712 (2015).
- 211 Fulda, S., Galluzzi, L. & Kroemer, G. Targeting mitochondria for cancer therapy. *Nat Rev Drug Discov* **9**, 447-464, doi:10.1038/nrd3137 (2010).
- 212 Bi, R., Li, W. L., Chen, M. Q., Zhu, Z. & Yao, Y. G. Rapid identification of mtDNA somatic mutations in gastric cancer tissues based on the mtDNA phylogeny. *Mutat Res* **709-710**, 15-20, doi:10.1016/j.mrfmmm.2011.02.016 (2011).

- 213 Xia, J., Guo, X., Yan, J. & Deng, K. The role of miR-148a in gastric cancer. *J Cancer Res Clin Oncol* **140**, 1451-1456, doi:10.1007/s00432-014-1649-8 (2014).
- 214 Shi, Z., Wei, Q. & She, J. MicroRNAs in gastric cancer metastasis. *Crit Rev Eukaryot Gene Expr* **24**, 39-53, doi:10.1615/critreveukaryotgeneexpr.2014007896 (2014).
- 215 Vandevenne, M., Delmarcelle, M. & Galleni, M. RNA Regulatory Networks as a Control of Stochasticity in Biological Systems. *Front Genet* **10**, 403, doi:10.3389/fgene.2019.00403 (2019).
- 216 Ahnert, S. E., Fink, T. M. & Zinovyev, A. How much non-coding DNA do eukaryotes require? *J Theor Biol* **252**, 587-592, doi:10.1016/j.jtbi.2008.02.005 (2008).
- 217 Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A. & Luscombe, N. M. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet* **10**, 252-263, doi:10.1038/nrg2538 (2009).
- 218 Li, J. & Liu, C. Coding or Noncoding, the Converging Concepts of RNAs. *Front Genet* **10**, 496, doi:10.3389/fgene.2019.00496 (2019).
- 219 Kellis, M. *et al.* Defining functional DNA elements in the human genome. *Proc Natl Acad Sci U S A* **111**, 6131-6138, doi:10.1073/pnas.1318948111 (2014).
- 220 Dahariya, S. *et al.* Long non-coding RNA: Classification, biogenesis and functions in blood cells. *Mol Immunol* **112**, 82-92, doi:10.1016/j.molimm.2019.04.011 (2019).
- 221 Karlsson, O. & Baccarelli, A. A. Environmental Health and Long Non-coding RNAs. *Curr Environ Health Rep* **3**, 178-187, doi:10.1007/s40572-016-0092-1 (2016).
- 222 Sun, Q., Hao, Q. & Prasanth, K. V. Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. *Trends Genet* **34**, 142-157, doi:10.1016/j.tig.2017.11.005 (2018).
- 223 Fatica, A. & Bozzoni, I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet* **15**, 7-21, doi:10.1038/nrg3606 (2014).
- 224 Ponting, C. P., Oliver, P. L. & Reik, W. Evolution and functions of long noncoding RNAs. *Cell* **136**, 629-641, doi:10.1016/j.cell.2009.02.006 (2009).
- 225 Uszczynska-Ratajczak, B., Lagarde, J., Frankish, A., Guigo, R. & Johnson, R. Towards a complete map of the human long non-coding RNA transcriptome. *Nat Rev Genet* **19**, 535-548, doi:10.1038/s41576-018-0017-y (2018).
- 226 Sun, B. *et al.* Research progress on the interactions between long non-coding RNAs and microRNAs in human cancer. *Oncol Lett* **19**, 595-605, doi:10.3892/ol.2019.11182 (2020).
- 227 Ard, R., Tong, P. & Allshire, R. C. Long non-coding RNA-mediated transcriptional interference of a permease gene confers drug tolerance in fission yeast. *Nat Commun* **5**, 5576, doi:10.1038/ncomms6576 (2014).
- 228 Zhang, X. *et al.* Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. *Int J Mol Sci* **20**, doi:10.3390/ijms20225573 (2019).
- 229 Li, X. & Fu, X. D. Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat Rev Genet* **20**, 503-519, doi:10.1038/s41576-019-0135-1 (2019).
- 230 Begolli, R., Sideris, N. & Giakountis, A. LncRNAs as Chromatin Regulators in Cancer: From Molecular Function to Clinical Potential. *Cancers (Basel)* **11**, doi:10.3390/cancers11101524 (2019).
- 231 Sarropoulos, I., Marin, R., Cardoso-Moreira, M. & Kaessmann, H. Developmental dynamics of lncRNAs across mammalian organs and species. *Nature* **571**, 510-514, doi:10.1038/s41586-019-1341-x (2019).
- 232 Alessio, E., Bonadio, R. S., Buson, L., Chemello, F. & Cagnin, S. A Single Cell but Many Different Transcripts: A Journey into the World of Long Non-Coding RNAs. *Int J Mol Sci* **21**, doi:10.3390/ijms21010302 (2020).
- 233 Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. *Annu Rev Biochem* **81**, 145-166, doi:10.1146/annurev-biochem-051410-092902 (2012).

- 234 Guttman, M. & Rinn, J. L. Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339-346, doi:10.1038/nature10887 (2012).
- 235 Batista, P. J. & Chang, H. Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298-1307, doi:10.1016/j.cell.2013.02.012 (2013).
- 236 Novikova, I. V., Hennelly, S. P. & Sanbonmatsu, K. Y. Sizing up long non-coding RNAs: do lncRNAs have secondary and tertiary structure? *Bioarchitecture* **2**, 189-199, doi:10.4161/bioa.22592 (2012).
- 237 Beltran, M. *et al.* A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* **22**, 756-769, doi:10.1101/gad.455708 (2008).
- 238 Tsai, M. C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689-693, doi:10.1126/science.1192002 (2010).
- 239 Willingham, A. T. *et al.* A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* **309**, 1570-1573, doi:10.1126/science.1115901 (2005).
- 240 Clemson, C. M. *et al.* An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33**, 717-726, doi:10.1016/j.molcel.2009.01.026 (2009).
- 241 Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322**, 750-756, doi:10.1126/science.1163045 (2008).
- 242 Rinn, J. L. *et al.* Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129**, 1311-1323, doi:10.1016/j.cell.2007.05.022 (2007).
- 243 Wang, J. *et al.* CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. *Nucleic Acids Res* **38**, 5366-5383, doi:10.1093/nar/gkq285 (2010).
- 244 Carrieri, C. *et al.* Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**, 454-457, doi:10.1038/nature11508 (2012).
- 245 Pang, K. C. *et al.* Genome-wide identification of long noncoding RNAs in CD8+ T cells. *J Immunol* **182**, 7738-7748, doi:10.4049/jimmunol.0900603 (2009).
- 246 Golden, D. E., Gerbasi, V. R. & Sontheimer, E. J. An inside job for siRNAs. *Mol Cell* **31**, 309-312, doi:10.1016/j.molcel.2008.07.008 (2008).
- 247 Cai, X. & Cullen, B. R. The imprinted H19 noncoding RNA is a primary microRNA precursor. *RNA* **13**, 313-316, doi:10.1261/rna.351707 (2007).
- 248 Slavoff, S. A. *et al.* Peptidomic discovery of short open reading frame-encoded peptides in human cells. *Nat Chem Biol* **9**, 59-64, doi:10.1038/nchembio.1120 (2013).
- 249 Ingolia, N. T., Lareau, L. F. & Weissman, J. S. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* **147**, 789-802, doi:10.1016/j.cell.2011.10.002 (2011).
- 250 Richard, J. L. C. & Eichhorn, P. J. A. Deciphering the roles of lncRNAs in breast development and disease. *Oncotarget* **9**, 20179-20212, doi:10.18632/oncotarget.24591 (2018).
- 251 Chen, L. L. Linking Long Noncoding RNA Localization and Function. *Trends Biochem Sci* **41**, 761-772, doi:10.1016/j.tibs.2016.07.003 (2016).
- 252 Chen, L. L. & Carmichael, G. G. Decoding the function of nuclear long non-coding RNAs. *Curr Opin Cell Biol* **22**, 357-364, doi:10.1016/j.ceb.2010.03.003 (2010).
- 253 Noh, J. H., Kim, K. M., McClusky, W. G., Abdelmohsen, K. & Gorospe, M. Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip Rev RNA* **9**, e1471, doi:10.1002/wrna.1471 (2018).

- 254 Fu, P. F., Zheng, X., Fan, X. & Lin, A. F. Role of cytoplasmic lncRNAs in regulating cancer  
255 signaling pathways. *J Zhejiang Univ Sci B* **20**, 1-8, doi:10.1631/jzus.B1800254 (2019).  
256 A focus on regulatory RNAs. *Nat Cell Biol* **21**, 535, doi:10.1038/s41556-019-0325-2  
(2019).
- 256 Kopp, F. & Mendell, J. T. Functional Classification and Experimental Dissection of Long  
257 Noncoding RNAs. *Cell* **172**, 393-407, doi:10.1016/j.cell.2018.01.011 (2018).
- 257 Miao, H. *et al.* A long noncoding RNA distributed in both nucleus and cytoplasm  
operates in the PYCARD-regulated apoptosis by coordinating the epigenetic and  
translational regulation. *PLoS Genet* **15**, e1008144,  
doi:10.1371/journal.pgen.1008144 (2019).
- 258 Cabili, M. N. *et al.* Localization and abundance analysis of human lncRNAs at single-  
cell and single-molecule resolution. *Genome Biol* **16**, 20, doi:10.1186/s13059-015-  
0586-4 (2015).
- 259 Gil, N. & Ulitsky, I. Regulation of gene expression by cis-acting long non-coding RNAs.  
*Nat Rev Genet* **21**, 102-117, doi:10.1038/s41576-019-0184-5 (2020).
- 260 Kuo, C. C. *et al.* Detection of RNA-DNA binding sites in long noncoding RNAs. *Nucleic  
Acids Res* **47**, e32, doi:10.1093/nar/gkz037 (2019).
- 261 Kim, T. K., Hemberg, M. & Gray, J. M. Enhancer RNAs: a class of long noncoding RNAs  
synthesized at enhancers. *Cold Spring Harb Perspect Biol* **7**, a018622,  
doi:10.1101/cshperspect.a018622 (2015).
- 262 Xiang, J. F. *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-  
range chromatin interactions at the MYC locus. *Cell Res* **24**, 513-531,  
doi:10.1038/cr.2014.35 (2014).
- 263 Liu, Z., Chen, Q. & Hann, S. S. The functions and oncogenic roles of CCAT1 in human  
cancer. *Biomed Pharmacother* **115**, 108943, doi:10.1016/j.biopha.2019.108943  
(2019).
- 264 Terranova, R. *et al.* Polycomb group proteins Ezh2 and Rnf2 direct genomic  
contraction and imprinted repression in early mouse embryos. *Dev Cell* **15**, 668-679,  
doi:10.1016/j.devcel.2008.08.015 (2008).
- 265 Pandey, R. R. *et al.* Kcnq1ot1 antisense noncoding RNA mediates lineage-specific  
transcriptional silencing through chromatin-level regulation. *Mol Cell* **32**, 232-246,  
doi:10.1016/j.molcel.2008.08.022 (2008).
- 266 Nagano, T. *et al.* The Air noncoding RNA epigenetically silences transcription by  
targeting G9a to chromatin. *Science* **322**, 1717-1720, doi:10.1126/science.1163802  
(2008).
- 267 Postepska-Igielska, A. *et al.* lncRNA Khps1 Regulates Expression of the Proto-  
oncogene SPHK1 via Triplex-Mediated Changes in Chromatin Structure. *Mol Cell* **60**,  
626-636, doi:10.1016/j.molcel.2015.10.001 (2015).
- 268 Hacisuleyman, E. *et al.* Topological organization of multichromosomal regions by the  
long intergenic noncoding RNA Firre. *Nat Struct Mol Biol* **21**, 198-206,  
doi:10.1038/nsmb.2764 (2014).
- 269 Zimta, A. A. *et al.* An Emerging Class of Long Non-coding RNA With Oncogenic Role  
Arises From the snoRNA Host Genes. *Front Oncol* **10**, 389,  
doi:10.3389/fonc.2020.00389 (2020).
- 270 Yin, Q. F. *et al.* Long noncoding RNAs with snoRNA ends. *Mol Cell* **48**, 219-230,  
doi:10.1016/j.molcel.2012.07.033 (2012).
- 271 Wang, X. *et al.* Induced ncRNAs allosterically modify RNA-binding proteins in cis to  
inhibit transcription. *Nature* **454**, 126-130, doi:10.1038/nature06992 (2008).
- 272 Sigova, A. A. *et al.* Transcription factor trapping by RNA in gene regulatory elements.  
*Science* **350**, 978-981, doi:10.1126/science.aad3346 (2015).

- 273 Yan, P., Luo, S., Lu, J. Y. & Shen, X. Cis- and trans-acting lncRNAs in pluripotency and reprogramming. *Curr Opin Genet Dev* **46**, 170-178, doi:10.1016/j.gde.2017.07.009 (2017).
- 274 Andersen, R. E. *et al.* The Long Noncoding RNA Pnky Is a Trans-acting Regulator of Cortical Development In Vivo. *Dev Cell* **49**, 632-642 e637, doi:10.1016/j.devcel.2019.04.032 (2019).
- 275 Yu, X. & Li, Z. Long non-coding RNA HOTAIR: A novel oncogene (Review). *Mol Med Rep* **12**, 5611-5618, doi:10.3892/mmr.2015.4161 (2015).
- 276 Qin, W. *et al.* Long non-coding RNA HOTAIR promotes tumorigenesis and forecasts a poor prognosis in cholangiocarcinoma. *Sci Rep* **8**, 12176, doi:10.1038/s41598-018-29737-4 (2018).
- 277 Hajjari, M. & Salavaty, A. HOTAIR: an oncogenic long non-coding RNA in different cancers. *Cancer Biol Med* **12**, 1-9, doi:10.7497/j.issn.2095-3941.2015.0006 (2015).
- 278 Gupta, R. A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071-1076, doi:10.1038/nature08975 (2010).
- 279 Chu, C., Qu, K., Zhong, F. L., Artandi, S. E. & Chang, H. Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* **44**, 667-678, doi:10.1016/j.molcel.2011.08.027 (2011).
- 280 Xing, Z. *et al.* lncRNA directs cooperative epigenetic regulation downstream of chemokine signals. *Cell* **159**, 1110-1125, doi:10.1016/j.cell.2014.10.013 (2014).
- 281 Tseng, Y. Y. *et al.* PVT1 dependence in cancer with MYC copy-number increase. *Nature* **512**, 82-86, doi:10.1038/nature13311 (2014).
- 282 Chujo, T. & Hirose, T. Nuclear Bodies Built on Architectural Long Noncoding RNAs: Unifying Principles of Their Construction and Function. *Mol Cells* **40**, 889-896, doi:10.14348/molcells.2017.0263 (2017).
- 283 Tripathi, V. *et al.* The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39**, 925-938, doi:10.1016/j.molcel.2010.08.011 (2010).
- 284 Sun, Y. & Ma, L. New Insights into Long Non-Coding RNA MALAT1 in Cancer and Metastasis. *Cancers (Basel)* **11**, doi:10.3390/cancers11020216 (2019).
- 285 Arun, G. *et al.* Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. *Genes Dev* **30**, 34-51, doi:10.1101/gad.270959.115 (2016).
- 286 Zhu, Z. *et al.* Knockdown long noncoding RNA nuclear paraspeckle assembly transcript 1 suppresses colorectal cancer through modulating miR-193a-3p/KRAS. *Cancer Med* **8**, 261-275, doi:10.1002/cam4.1798 (2019).
- 287 Sunwoo, H. *et al.* MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* **19**, 347-359, doi:10.1101/gr.087775.108 (2009).
- 288 Chen, L. L. & Carmichael, G. G. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell* **35**, 467-478, doi:10.1016/j.molcel.2009.06.027 (2009).
- 289 West, J. A. *et al.* The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell* **55**, 791-802, doi:10.1016/j.molcel.2014.07.012 (2014).
- 290 Naganuma, T. *et al.* Alternative 3'-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J* **31**, 4020-4034, doi:10.1038/emboj.2012.251 (2012).
- 291 Liu, B. *et al.* A cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkkappaB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* **27**, 370-381, doi:10.1016/j.ccell.2015.02.004 (2015).

- 292 Lin, A. *et al.* The LINK-A lncRNA activates normoxic HIF1 $\alpha$  signalling in triple-  
negative breast cancer. *Nat Cell Biol* **18**, 213-224, doi:10.1038/ncb3295 (2016).
- 293 Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering  
PUMILIO Proteins. *Cell* **164**, 69-80, doi:10.1016/j.cell.2015.12.017 (2016).
- 294 Cesana, M. *et al.* A long noncoding RNA controls muscle differentiation by functioning  
as a competing endogenous RNA. *Cell* **147**, 358-369, doi:10.1016/j.cell.2011.09.028  
(2011).
- 295 Quinn, J. J. & Chang, H. Y. Unique features of long non-coding RNA biogenesis and  
function. *Nat Rev Genet* **17**, 47-62, doi:10.1038/nrg.2015.10 (2016).
- 296 Wang, D. Q. *et al.* Long Non-coding RNAs, Novel Culprits, or Bodyguards in  
Neurodegenerative Diseases. *Mol Ther Nucleic Acids* **10**, 269-276,  
doi:10.1016/j.omtn.2017.12.011 (2018).
- 297 Nelson, B. R. *et al.* A peptide encoded by a transcript annotated as long noncoding  
RNA enhances SERCA activity in muscle. *Science* **351**, 271-275,  
doi:10.1126/science.aad4076 (2016).
- 298 Laressergues, D. *et al.* Primary transcripts of microRNAs encode regulatory peptides.  
*Nature* **520**, 90-93, doi:10.1038/nature14346 (2015).
- 299 Cohen, S. M. Everything old is new again: (linc)RNAs make proteins! *EMBO J* **33**, 937-  
938, doi:10.1002/embj.201488303 (2014).
- 300 Anderson, D. M. *et al.* A micropeptide encoded by a putative long noncoding RNA  
regulates muscle performance. *Cell* **160**, 595-606, doi:10.1016/j.cell.2015.01.009  
(2015).
- 301 Cheetham, S. W., Gruhl, F., Mattick, J. S. & Dinger, M. E. Long noncoding RNAs and the  
genetics of cancer. *Br J Cancer* **108**, 2419-2425, doi:10.1038/bjc.2013.233 (2013).
- 302 Khurana, E. *et al.* Role of non-coding sequence variants in cancer. *Nat Rev Genet* **17**,  
93-108, doi:10.1038/nrg.2015.17 (2016).
- 303 Wei, L. *et al.* Noncoding RNAs in gastric cancer: implications for drug resistance. *Mol  
Cancer* **19**, 62, doi:10.1186/s12943-020-01185-7 (2020).
- 304 Xu, Z. *et al.* The Association of HOTAIR with the Diagnosis and Prognosis of Gastric  
Cancer and Its Effect on the Proliferation of Gastric Cancer Cells. *Can J Gastroenterol  
Hepatol* **2019**, 3076345, doi:10.1155/2019/3076345 (2019).
- 305 Xu, Z. Y. *et al.* Knockdown of long non-coding RNA HOTAIR suppresses tumor invasion  
and reverses epithelial-mesenchymal transition in gastric cancer. *Int J Biol Sci* **9**, 587-  
597, doi:10.7150/ijbs.6339 (2013).
- 306 Zhou, X., Yin, C., Dang, Y., Ye, F. & Zhang, G. Identification of the long non-coding RNA  
H19 in plasma as a novel biomarker for diagnosis of gastric cancer. *Sci Rep* **5**, 11516,  
doi:10.1038/srep11516 (2015).
- 307 Liu, Y. *et al.* lncRNA GAS5 enhances G1 cell cycle arrest via binding to YBX1 to regulate  
p21 expression in stomach cancer. *Sci Rep* **5**, 10159, doi:10.1038/srep10159 (2015).
- 308 Sun, M. *et al.* Downregulated long noncoding RNA MEG3 is associated with poor  
prognosis and promotes cell proliferation in gastric cancer. *Tumour Biol* **35**, 1065-  
1073, doi:10.1007/s13277-013-1142-z (2014).
- 309 Wu, M. *et al.* Comprehensive Analysis of the Expression Profiles of Long Non-Coding  
RNAs with Associated ceRNA Network Involved in the Colon Cancer Staging and  
Progression. *Sci Rep* **9**, 16910, doi:10.1038/s41598-019-52883-2 (2019).
- 310 Li, T., Mo, X., Fu, L., Xiao, B. & Guo, J. Molecular mechanisms of long noncoding RNAs  
on gastric cancer. *Oncotarget* **7**, 8601-8612, doi:10.18632/oncotarget.6926 (2016).
- 311 Gu, Y. *et al.* lncRNAs: emerging biomarkers in gastric cancer. *Future Oncol* **11**, 2427-  
2441, doi:10.2217/fon.15.175 (2015).



- 312 Machlowska, J., Baj, J., Sitarz, M., Maciejewski, R. & Sitarz, R. Gastric Cancer: Epidemiology, Risk Factors, Classification, Genomic Characteristics and Treatment Strategies. *Int J Mol Sci* **21**, doi:10.3390/ijms21114012 (2020).
- 313 Orditura, M. *et al.* Treatment of gastric cancer. *World J Gastroenterol* **20**, 1635-1649, doi:10.3748/wjg.v20.i7.1635 (2014).
- 314 Riihimaki, M., Hemminki, A., Sundquist, K., Sundquist, J. & Hemminki, K. Metastatic spread in patients with gastric cancer. *Oncotarget* **7**, 52307-52316, doi:10.18632/oncotarget.10740 (2016).
- 315 Wagner, A. D. *et al.* Chemotherapy for advanced gastric cancer. *Cochrane Database Syst Rev* **8**, CD004064, doi:10.1002/14651858.CD004064.pub4 (2017).
- 316 Matsuoka, T. & Yashiro, M. Biomarkers of gastric cancer: Current topics and future perspective. *World J Gastroenterol* **24**, 2818-2832, doi:10.3748/wjg.v24.i26.2818 (2018).
- 317 Colquhoun, A. *et al.* Global patterns of cardia and non-cardia gastric cancer incidence in 2012. *Gut* **64**, 1881-1888, doi:10.1136/gutjnl-2014-308915 (2015).
- 318 Maleki, S. S. & Rocken, C. Chromosomal Instability in Gastric Cancer Biology. *Neoplasia* **19**, 412-420, doi:10.1016/j.neo.2017.02.012 (2017).
- 319 Chen, L., Zhou, W., Zhang, L. & Zhang, F. Genome architecture and its roles in human copy number variation. *Genomics Inform* **12**, 136-144, doi:10.5808/GI.2014.12.4.136 (2014).
- 320 Kwon, M. J. *et al.* Genes co-amplified with ERBB2 or MET as novel potential cancer-promoting genes in gastric cancer. *Oncotarget* **8**, 92209-92226, doi:10.18632/oncotarget.21150 (2017).
- 321 Yang, Y. *et al.* Recurrently deregulated lncRNAs in hepatocellular carcinoma. *Nat Commun* **8**, 14421, doi:10.1038/ncomms14421 (2017).
- 322 Liu, H. *et al.* Copy number variations primed lncRNAs deregulation contribute to poor prognosis in colorectal cancer. *Aging (Albany NY)* **11**, 6089-6108, doi:10.18632/aging.102168 (2019).
- 323 Zheng, M. *et al.* Identification three lncRNA prognostic signature of ovarian cancer based on genome-wide copy number variation. *Biomed Pharmacother* **124**, 109810, doi:10.1016/j.biopha.2019.109810 (2020).
- 324 et.al., H. Long Noncoding RNA GAPLINC Regulates CD44-Dependent Cell Invasiveness and Associates with Poor Prognosis of Gastric Cancer. *Cancer Research* **74**, doi:doi: 10.1158/0008-5472.CAN-14-0686 (2014).
- 325 Du, W. *et al.* Identification of prognostic biomarkers of hepatocellular carcinoma via long noncoding RNAs expression and copy number alterations. *Epigenomics*, doi:10.2217/epi-2019-0385 (2020).
- 326 Hastings, P. J., Lupski, J. R., Rosenberg, S. M. & Ira, G. Mechanisms of change in gene copy number. *Nat Rev Genet* **10**, 551-564, doi:10.1038/nrg2593 (2009).
- 327 de Koning, A. P., Gu, W., Castoe, T. A., Batzer, M. A. & Pollock, D. D. Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* **7**, e1002384, doi:10.1371/journal.pgen.1002384 (2011).
- 328 Kim, P. M. *et al.* Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. *Genome Res* **18**, 1865-1874, doi:10.1101/gr.081422.108 (2008).
- 329 Bailey, J. A. *et al.* Recent segmental duplications in the human genome. *Science* **297**, 1003-1007, doi:10.1126/science.1072047 (2002).
- 330 Park, S. S. *et al.* Structure and evolution of the Smith-Magenis syndrome repeat gene clusters, SMS-REPs. *Genome Res* **12**, 729-738, doi:10.1101/gr.82802 (2002).

- 331 Kurotaki, N., Stankiewicz, P., Wakui, K., Niikawa, N. & Lupski, J. R. Sotos syndrome common deletion is mediated by directly oriented subunits within inverted Sos-REP low-copy repeats. *Hum Mol Genet* **14**, 535-542, doi:10.1093/hmg/ddi050 (2005).
- 332 Linardopoulou, E. V. *et al.* Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* **437**, 94-100, doi:10.1038/nature04029 (2005).
- 333 Jiang, Z. *et al.* Ancestral reconstruction of segmental duplications reveals punctuated cores of human genome evolution. *Nat Genet* **39**, 1361-1368, doi:10.1038/ng.2007.9 (2007).
- 334 Dittwald, P. *et al.* NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits. *Genome Res* **23**, 1395-1409, doi:10.1101/gr.152454.112 (2013).
- 335 Pascual, R. N. R. a. J. M. Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric Disease. *Academic Press*, 1464, doi:<https://doi.org/10.1016/C2012-0-02688-9> (2015).
- 336 Arlt, M. F., Wilson, T. E. & Glover, T. W. Replication stress and mechanisms of CNV formation. *Curr Opin Genet Dev* **22**, 204-210, doi:10.1016/j.gde.2012.01.009 (2012).
- 337 Sharp, A. J. *et al.* Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* **38**, 1038-1042, doi:10.1038/ng1862 (2006).
- 338 Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B. & Hogness, D. S. Molecular genetics of inherited variation in human color vision. *Science* **232**, 203-210, doi:10.1126/science.3485310 (1986).
- 339 Urh, K. & Kunej, T. Genome-wide screening for smallest regions of overlaps in cryptorchidism. *Reprod Biomed Online* **37**, 85-99, doi:10.1016/j.rbmo.2018.02.008 (2018).
- 340 Soler-Alfonso, C. *et al.* CHRNA7 triplication associated with cognitive impairment and neuropsychiatric phenotypes in a three-generation pedigree. *Eur J Hum Genet* **22**, 1071-1076, doi:10.1038/ejhg.2013.302 (2014).
- 341 Ishmukhametova, A. *et al.* Dissecting the structure and mechanism of a complex duplication-triplication rearrangement in the DMD gene. *Hum Mutat* **34**, 1080-1084, doi:10.1002/humu.22353 (2013).
- 342 Carvalho, C. M. *et al.* Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nat Genet* **43**, 1074-1081, doi:10.1038/ng.944 (2011).
- 343 Beri, S., Bonaglia, M. C. & Giorda, R. Low-copy repeats at the human VIPR2 gene predispose to recurrent and nonrecurrent rearrangements. *Eur J Hum Genet* **21**, 757-761, doi:10.1038/ejhg.2012.235 (2013).
- 344 Beck, C. R. *et al.* Complex genomic rearrangements at the PLP1 locus include triplication and quadruplication. *PLoS Genet* **11**, e1005050, doi:10.1371/journal.pgen.1005050 (2015).
- 345 Gu, S. *et al.* Alu-mediated diverse and complex pathogenic copy-number variants within human chromosome 17 at p13.3. *Hum Mol Genet* **24**, 4061-4077, doi:10.1093/hmg/ddv146 (2015).
- 346 Wu, Z. *et al.* Equilibrating strand exchange promoted by DNA conformational switching. *Sci Rep* **3**, 1121, doi:10.1038/srep01121 (2013).
- 347 Carvalho, C. M. *et al.* Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching. *Hum Mol Genet* **18**, 2188-2203, doi:10.1093/hmg/ddp151 (2009).

- 348 Liu, P. *et al.* Frequency of nonallelic homologous recombination is correlated with length of homology: evidence that ectopic synapsis precedes ectopic crossing-over. *Am J Hum Genet* **89**, 580-588, doi:10.1016/j.ajhg.2011.09.009 (2011).
- 349 Kim, P. M., Paffett, K. S., Solinger, J. A., Heyer, W. D. & Nickoloff, J. A. Spontaneous and double-strand break-induced recombination, and gene conversion tract lengths, are differentially affected by overexpression of wild-type or ATPase-defective yeast Rad54. *Nucleic Acids Res* **30**, 2727-2735, doi:10.1093/nar/gkf413 (2002).
- 350 Small, K. & Warren, S. T. Emerin deletions occurring on both Xq28 inversion backgrounds. *Hum Mol Genet* **7**, 135-139, doi:10.1093/hmg/7.1.135 (1998).
- 351 Inoue, K. *et al.* Genomic rearrangements resulting in PLP1 deletion occur by nonhomologous end joining and cause different dysmyelinating phenotypes in males and females. *Am J Hum Genet* **71**, 838-853, doi:10.1086/342728 (2002).
- 352 Stankiewicz, P. & Lupski, J. R. Genome architecture, rearrangements and genomic disorders. *Trends Genet* **18**, 74-82, doi:10.1016/s0168-9525(02)02592-1 (2002).
- 353 Liu, C. *et al.* A fine-scale dissection of the DNA double-strand break repair machinery and its implications for breast cancer therapy. *Nucleic Acids Res* **42**, 6106-6127, doi:10.1093/nar/gku284 (2014).
- 354 Liu, P., Carvalho, C. M., Hastings, P. J. & Lupski, J. R. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* **22**, 211-220, doi:10.1016/j.gde.2012.02.012 (2012).
- 355 Woodward, K. J. *et al.* Heterogeneous duplications in patients with Pelizaeus-Merzbacher disease suggest a mechanism of coupled homologous and nonhomologous recombination. *Am J Hum Genet* **77**, 966-987, doi:10.1086/498048 (2005).
- 356 Boone, P. M. *et al.* The Alu-rich genomic architecture of SPAST predisposes to diverse and functionally distinct disease-associated CNV alleles. *Am J Hum Genet* **95**, 143-161, doi:10.1016/j.ajhg.2014.06.014 (2014).
- 357 Vissers, L. E. *et al.* Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. *Hum Mol Genet* **18**, 3579-3593, doi:10.1093/hmg/ddp306 (2009).
- 358 Stankiewicz, P. *et al.* Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. *Am J Hum Genet* **84**, 780-791, doi:10.1016/j.ajhg.2009.05.005 (2009).
- 359 Vissers, L. E. & Stankiewicz, P. Microdeletion and microduplication syndromes. *Methods Mol Biol* **838**, 29-75, doi:10.1007/978-1-61779-507-7\_2 (2012).
- 360 Turner, D. J. *et al.* Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet* **40**, 90-95, doi:10.1038/ng.2007.40 (2008).
- 361 Robberecht, C., Voet, T., Zamani Esteki, M., Nowakowska, B. A. & Vermeesch, J. R. Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. *Genome Res* **23**, 411-418, doi:10.1101/gr.145631.112 (2013).
- 362 Ou, Z. *et al.* Observation and prediction of recurrent human translocations mediated by NAHR between nonhomologous chromosomes. *Genome Res* **21**, 33-46, doi:10.1101/gr.111609.110 (2011).
- 363 Golzio, C. & Katsanis, N. Genetic architecture of reciprocal CNVs. *Curr Opin Genet Dev* **23**, 240-248, doi:10.1016/j.gde.2013.04.013 (2013).
- 364 Dittwald, P. *et al.* Inverted low-copy repeats and genome instability--a genome-wide analysis. *Hum Mutat* **34**, 210-220, doi:10.1002/humu.22217 (2013).

- 365 Liu, P. *et al.* Mechanism, prevalence, and more severe neuropathy phenotype of the Charcot-Marie-Tooth type 1A triplication. *Am J Hum Genet* **94**, 462-469, doi:10.1016/j.ajhg.2014.01.017 (2014).
- 366 Berg, I. L. *et al.* PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat Genet* **42**, 859-863, doi:10.1038/ng.658 (2010).
- 367 Zhang, F. *et al.* Identification of uncommon recurrent Potocki-Lupski syndrome-associated duplications and the distribution of rearrangement types and mechanisms in PTLs. *Am J Hum Genet* **86**, 462-470, doi:10.1016/j.ajhg.2010.02.001 (2010).
- 368 Cooper, G. M. *et al.* A copy number variation morbidity map of developmental delay. *Nat Genet* **43**, 838-846, doi:10.1038/ng.909 (2011).
- 369 Konstantinopoulos, P. A., Ceccaldi, R., Shapiro, G. I. & D'Andrea, A. D. Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. *Cancer Discov* **5**, 1137-1154, doi:10.1158/2159-8290.CD-15-0714 (2015).
- 370 Nazaryan-Petersen, L. *et al.* Replicative and non-replicative mechanisms in the formation of clustered CNVs are indicated by whole genome characterization. *PLoS Genet* **14**, e1007780, doi:10.1371/journal.pgen.1007780 (2018).
- 371 Grajcarek, J. *et al.* Genome-wide microhomologies enable precise template-free editing of biologically relevant deletion mutations. *Nat Commun* **10**, 4856, doi:10.1038/s41467-019-12829-8 (2019).
- 372 Hsiao, M. C. *et al.* Decoding NF1 Intragenic Copy-Number Variations. *Am J Hum Genet* **97**, 238-249, doi:10.1016/j.ajhg.2015.06.002 (2015).
- 373 Callinan, P. A. & Batzer, M. A. Retrotransposable elements and human disease. *Genome Dyn* **1**, 104-115, doi:10.1159/000092503 (2006).
- 374 Boone, P. M. *et al.* Alu-specific microhomology-mediated deletion of the final exon of SPAST in three unrelated subjects with hereditary spastic paraplegia. *Genet Med* **13**, 582-592, doi:10.1097/GIM.0b013e3182106775 (2011).
- 375 Stankiewicz, P., Pursley, A. N. & Cheung, S. W. Challenges in clinical interpretation of microduplications detected by array CGH analysis. *Am J Med Genet A* **152A**, 1089-1100, doi:10.1002/ajmg.a.33216 (2010).
- 376 Deininger, P. L. & Batzer, M. A. Alu repeats and human disease. *Mol Genet Metab* **67**, 183-193, doi:10.1006/mgme.1999.2864 (1999).
- 377 Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol* **20**, 698-714, doi:10.1038/s41580-019-0152-0 (2019).
- 378 Ottaviani, D., LeCain, M. & Sheer, D. The role of microhomology in genomic structural variation. *Trends Genet* **30**, 85-94, doi:10.1016/j.tig.2014.01.001 (2014).
- 379 Cho, N. W. & Greenberg, R. A. DNA repair: Familiar ends with alternative endings. *Nature* **518**, 174-176, doi:10.1038/nature14200 (2015).
- 380 Lee, J. A., Carvalho, C. M. & Lupski, J. R. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell* **131**, 1235-1247, doi:10.1016/j.cell.2007.11.037 (2007).
- 381 Tanaka, H. & Yao, M. C. Palindromic gene amplification--an evolutionarily conserved role for DNA inverted repeats in the genome. *Nat Rev Cancer* **9**, 216-224, doi:10.1038/nrc2591 (2009).
- 382 Tanaka, H., Bergstrom, D. A., Yao, M. C. & Tapscott, S. J. Large DNA palindromes as a common form of structural chromosome aberrations in human cancers. *Hum Cell* **19**, 17-23, doi:10.1111/j.1749-0774.2005.00003.x (2006).
- 383 Schmidt, K. H., Wu, J. & Kolodner, R. D. Control of translocations between highly diverged genes by Sgs1, the *Saccharomyces cerevisiae* homolog of the Bloom's syndrome protein. *Mol Cell Biol* **26**, 5406-5420, doi:10.1128/MCB.00161-06 (2006).

- 384 Hastings, P. J., Ira, G. & Lupski, J. R. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet* **5**, e1000327, doi:10.1371/journal.pgen.1000327 (2009).
- 385 Shaw, C. J. & Lupski, J. R. Non-recurrent 17p11.2 deletions are generated by homologous and non-homologous mechanisms. *Hum Genet* **116**, 1-7, doi:10.1007/s00439-004-1204-9 (2005).
- 386 Lee, J. A. *et al.* Role of genomic architecture in PLP1 duplication causing Pelizaeus-Merzbacher disease. *Hum Mol Genet* **15**, 2250-2265, doi:10.1093/hmg/ddl150 (2006).
- 387 Sheen, C. R. *et al.* Double complex mutations involving F8 and FUNDC2 caused by distinct break-induced replication. *Hum Mutat* **28**, 1198-1206, doi:10.1002/humu.20591 (2007).
- 388 Payen, C., Koszul, R., Dujon, B. & Fischer, G. Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet* **4**, e1000175, doi:10.1371/journal.pgen.1000175 (2008).
- 389 Deem, A. *et al.* Defective break-induced replication leads to half-crossovers in *Saccharomyces cerevisiae*. *Genetics* **179**, 1845-1860, doi:10.1534/genetics.108.087940 (2008).
- 390 Bauters, M. *et al.* Nonrecurrent MECP2 duplications mediated by genomic architecture-driven DNA breaks and break-induced replication repair. *Genome Res* **18**, 847-858, doi:10.1101/gr.075903.107 (2008).
- 391 Kuo, M. T., Vyas, R. C., Jiang, L. X. & Hittelman, W. N. Chromosome breakage at a major fragile site associated with P-glycoprotein gene amplification in multidrug-resistant CHO cells. *Mol Cell Biol* **14**, 5202-5211, doi:10.1128/mcb.14.8.5202 (1994).
- 392 Durkin, S. G. *et al.* Replication stress induces tumor-like microdeletions in FHIT/FRA3B. *Proc Natl Acad Sci U S A* **105**, 246-251, doi:10.1073/pnas.0708097105 (2008).
- 393 Coquelle, A., Rozier, L., Dutrillaux, B. & Debatisse, M. Induction of multiple double-strand breaks within an hsr by meganuclease-Scel expression or fragile site activation leads to formation of double minutes and other chromosomal rearrangements. *Oncogene* **21**, 7671-7679, doi:10.1038/sj.onc.1205880 (2002).
- 394 Coquelle, A., Pipiras, E., Toledo, F., Buttin, G. & Debatisse, M. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. *Cell* **89**, 215-225, doi:10.1016/s0092-8674(00)80201-9 (1997).
- 395 Arlt, M. F. *et al.* Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants. *Am J Hum Genet* **84**, 339-350, doi:10.1016/j.ajhg.2009.01.024 (2009).
- 396 Michel, B., Ehrlich, S. D. & Uzest, M. DNA double-strand breaks caused by replication arrest. *EMBO J* **16**, 430-438, doi:10.1093/emboj/16.2.430 (1997).
- 397 Albertini, A. M., Hofer, M., Calos, M. P. & Miller, J. H. On the formation of spontaneous deletions: the importance of short sequence homologies in the generation of large deletions. *Cell* **29**, 319-328, doi:10.1016/0092-8674(82)90148-9 (1982).
- 398 Bi, X. & Liu, L. F. recA-independent and recA-dependent intramolecular plasmid recombination. Differential homology requirement and distance effect. *J Mol Biol* **235**, 414-423, doi:10.1006/jmbi.1994.1002 (1994).
- 399 Saveson, C. J. & Lovett, S. T. Enhanced deletion formation by aberrant DNA replication in *Escherichia coli*. *Genetics* **146**, 457-470 (1997).
- 400 Russo, A. PRINS evaluation of chromosome instability in mammalian cells by detection of repetitive DNA sequences in micronuclei. *Methods Mol Biol* **334**, 89-104, doi:10.1385/1-59745-068-5:89 (2006).
- 401 Bzymek, M. & Lovett, S. T. Instability of repetitive DNA sequences: the role of replication in multiple mechanisms. *Proc Natl Acad Sci U S A* **98**, 8319-8325, doi:10.1073/pnas.111008398 (2001).

- 402 Bierne, H., Vilette, D., Ehrlich, S. D. & Michel, B. Isolation of a dnaE mutation which  
enhances RecA-independent homologous recombination in the Escherichia coli  
chromosome. *Mol Microbiol* **24**, 1225-1234, doi:10.1046/j.1365-  
2958.1997.4381795.x (1997).
- 403 Spies, M. & Fishel, R. Mismatch repair during homologous and homeologous  
recombination. *Cold Spring Harb Perspect Biol* **7**, a022657,  
doi:10.1101/cshperspect.a022657 (2015).
- 404 Slack, A., Thornton, P. C., Magner, D. B., Rosenberg, S. M. & Hastings, P. J. On the  
mechanism of gene amplification induced under stress in Escherichia coli. *PLoS Genet*  
**2**, e48, doi:10.1371/journal.pgen.0020048 (2006).
- 405 Kugelberg, E., Kofoed, E., Reams, A. B., Andersson, D. I. & Roth, J. R. Multiple pathways  
of selected gene amplification during adaptive mutation. *Proc Natl Acad Sci U S A* **103**,  
17319-17324, doi:10.1073/pnas.0608309103 (2006).
- 406 Zhang, F. *et al.* The DNA replication FoSTeS/MMBIR mechanism can generate  
genomic, genic and exonic complex rearrangements in humans. *Nat Genet* **41**, 849-  
853, doi:10.1038/ng.399 (2009).
- 407 Zhang, F., Carvalho, C. M. & Lupski, J. R. Complex human chromosomal and genomic  
rearrangements. *Trends Genet* **25**, 298-307, doi:10.1016/j.tig.2009.05.005 (2009).