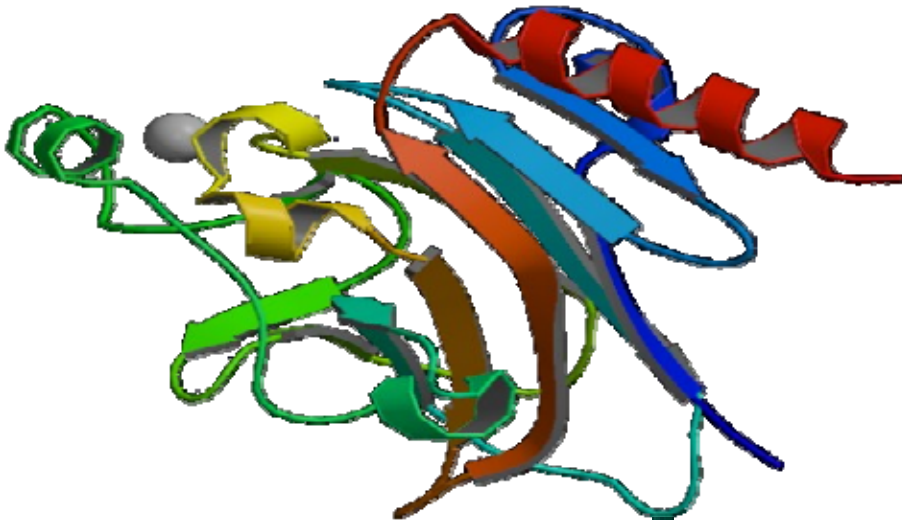


# ΑΝΑΣΤΑΣΙΟΣ ΚΥΒΕΝΤΙΔΗΣ

Μεταπτυχιακός Φοιτητής Ιατρικής Σχολής  
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

## ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Ανάλυση των Βάσεων Δεδομένων των  
Γενετικών και Σωματικών Μεταλλάξεων του  
Γονιδίου *p53*. Συσχέτιση με Πιθανούς  
Μηχανισμούς Μεθυλίωσης και  
Μεταλλαξιγένεσης.**



Οκτώβριος 2006

**Τριμελής επιτροπή:**

**Επιβλέπουσα: ΣΟΦΙΑ ΚΟΥΙΔΟΥ – ΑΝΔΡΕΟΥ**  
Αναπληρώτρια Καθηγήτρια Βιοχημείας Ιατρικής Σχολής  
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

**ΓΕΩΡΓΙΟΣ ΤΖΗΜΑΓΙΩΡΓΗΣ**  
Λέκτορας Βιοχημείας Ιατρικής Σχολής  
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

**ΝΙΚΟΛΑΟΣ ΜΑΓΚΛΑΒΕΡΑΣ**  
Αναπληρωτής Καθηγητής Πληροφορικής Ιατρικής Σχολής  
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

## ΠΕΡΙΛΗΨΗ

Μεγάλη πρόκληση της σημερινής εποχής είναι η ερμηνεία των μηχανισμών επιγενετικής τροποποίησης και μεταλλάξεων της αλληλουχίας των νουκλεοτιδίων στο μόριο του DNA. Στη νεότερη βιβλιογραφία αναφέρεται ότι η επιγενετική μεθυλίωση της κυτοσίνης είναι ικανή να αυξήσει τη μεταλλαξιμότητά της. Επίσης αναφέρεται ότι ορισμένα πρότυπα αλληλουχίας νουκλεοτιδίων είναι επιρρεπή σε μεταλλάξεις λόγω των εναλλακτικών τεταρτοταγών δομών που μπορούν να πάρουν στο χώρο. Στην παρούσα εργασία γίνεται ανάλυση γενετικών και σωματικών βάσεων μεταλλάξεων του ογκοκατασταλτικού γονιδίου *p53*. Οι μεταλλάξεις του *p53* αποτελούν αντικείμενο έντονης έρευνας και υπάρχουν 1.424 καταγραφές στη γενετική βάση δεδομένων και 21.587 καταγραφές στη σωματική. Υπάρχει, επομένως, μεγάλος όγκος πληροφοριών ώστε να μας οδηγήσει σε αποτελέσματα.

Η ανάλυση πραγματοποιήθηκε με τη βοήθεια: (α) προγράμματος διαχείρισης spreadsheet με παράλληλη στατιστική εκτίμηση των αποτελεσμάτων και (β) προγράμματος εξειδικευμένου στην ανάλυση φασμάτων μεταλλάξεων των γονιδίων (iMARS). Αναλύθηκε, με τα προγράμματα αυτά, η μεταλλαξιμότητα των κωδικίων, των δινουκλεοτιδίων και των νουκλεοτιδίων του *p53* στις γενετικές και σωματικές μεταλλάξεις. Έγινε ανάλυση των μεταλλάξεων σε σχέση με τις βάσεις που μεταλλάσσονται και το αποτέλεσμα της μετάλλαξης που προκαλούν. Επίσης, έγινε προσπάθεια αναγνώρισης των θέσεων υψηλής μεταλλαξιμότητας στις γενετικές και σωματικές μεταλλάξεις και για πρώτη φορά στη βιβλιογραφία επιχειρείται ερμηνεία τους με βάση την τριτοταγή και τεταρτοταγή διαμόρφωση που παίρνουν στο χώρο. Για πρώτη φορά επίσης γίνεται ανάλυση του συνόλου των γενετικών και σωματικών μεταλλάξεων του *p53* ανά νουκλεοτίδιο. Ακόμη, οι βάσεις δεδομένων αναλύθηκαν σύμφωνα με τύπους μεταλλάξεων που πιθανόν σχετίζονται με μεθυλίωση. Έγινε εκτίμηση της μεταλλαξιμότητας ανά αμινοξύ σύμφωνα με τα κωδικία που αντιστοιχούν σε αυτό. Παράλληλα, έγινε προσπάθεια αποκάλυψης της διεισδυτικότητας των γενετικών μεταλλάξεων και ανάλυση του φάσματος των μεταλλάξεων των γενετικών συνδρόμων ανάλογα με τα είδη καρκίνου που προκύπτουν. Επίσης μελετήθηκε η συμβολή της χρήσης καπνού και αλκοόλ στην αύξηση της καρκινογένεσης.

Από τα αποτελέσματα της εργασίας διαπιστώθηκε ότι ενώ όλες οι θέσεις CpG του γονιδίου είναι μεθυλιωμένες, ορισμένες μόνο εμφανίζουν υψηλή μεταλλαξιμότητα. Επίσης, προκύπτει ότι η αργινίνη, ως βασικό αμινοξύ και παράγοντας προσκόλλησης στο DNA, μεταλλάσσεται συντριπτικά περισσότερο από τα υπόλοιπα αμινοξέα και ότι τα γενετικά σύνδρομα φαίνεται να σχετίζονται με την παρουσία συγκεκριμένων μεταλλάξεων του γονιδίου.

Επιπρόσθετα, με τις ήδη διατυπωμένες θεωρίες δεν μπορεί να ερμηνευθεί ο χαμηλός μέσος όρος ηλικίας εμφάνισης καρκίνου των σιωπηρών σωματικών μεταλλάξεων και των μεταλλάξεων των ενδονίων των γενετικών κυττάρων. Από τις παρατηρήσεις που προαναφέρονται προκύπτει ότι οι παράγοντες επιλογής των μεταλλάξεων του *p53* των καρκινικών κυττάρων δεν περιορίζονται μόνο στην τοξικότητα του μεταλλαγμένου αμινοξέως ούτε στην παρουσία μεθυλίωσης, αλλά υποδηλώνουν την παρουσία επιπλέον παραγόντων που συμμετέχουν στην καρκινογένεση. Αυτοί οι παράγοντες πιθανόν είναι: (α) η παρουσία ασύμμετρης μεθυλίωσης της κυτοσίνης (δηλαδή εκτός θέσεων CpG) (β) η ανεπάρκεια επιδιορθωτικών πρωτεϊνών του DNA (γ) η αλλοίωση των θέσεων ματίσματος και ελαττωματικό μάτισμα και (δ) η δυνατότητα συγκεκριμένων αλληλουχιών νουκλεοτιδίων να παίρνουν εναλλακτικές διαμορφώσεις στο μόριο του DNA. Η ανάλυση των γενετικών και σωματικών βάσεων δεδομένων μπορεί να οδηγήσει στη διαμόρφωση χρήσιμων μοντέλων για τη μελέτη των παραπάνω παραγόντων και της καρκινογένεσης γενικότερα.

## ΠΕΡΙΕΧΟΜΕΝΑ

### ΠΕΡΙΛΗΨΗ

#### I. ΕΙΣΑΓΩΓΗ

- I.α. Η μεθυλίωση της κυτοσίνης
- I.β. Γενικά για την CpG μεθυλίωση
- I.γ. Οι μεθυλοτρανσφεράσες
- I.δ. Το γονίδιο *p53*
- I.ε. Τι πρότυπα νουκλεοτιδίων μελετούμε και για ποιο λόγο
- I.στ. Τα δινουκλεοτίδια TG και CA αποτελούν θέσεις αυξημένης γενετικής αστάθειας
- I.ζ. Ιδιότητες των τρινουκλεοτιδίων για σχηματισμό non-B DNA δομών
- I.η. Non B – DNA διαμορφώσεις
- I.θ. Σιωπηρές γενετικές μεταλλάξεις του *p53*
- I.ι. Γενετικά σύνδρομα που οφείλονται σε μεταλλάξεις του *p53*

#### II. ΥΛΙΚΟ – ΜΕΘΟΔΟΙ

- II.α. Προγράμματα ανάλυσης των βάσεων δεδομένων
- II.β. Στατιστική επεξεργασία

#### III. ΑΠΟΤΕΛΕΣΜΑΤΑ

##### 1. Χρήση προγράμματος *iMARS*

- III.1.α. Ποσοστά των τεσσάρων νουκλεοτιδίων στο γονίδιο του *p53*
- III.1.β. Μεταλλάξεις των τεσσάρων νουκλεοτιδίων.
- III.1.γ. Από τα ζεύγη νουκλεοτιδίων G:C και A:T διερευνάται ποιο νουκλεοτίδιο μεταλλάσσεται περισσότερο σε κάθε τύπο μετάλλαξης.
- III.1.δ. Συχνότητα τύπου μετάλλαξης κάθε δινουκλεοτιδίου
- III.1.ε. Κατανομή των γενετικών και σωματικών μεταλλάξεων ανά δινουκλεοτίδιο
- III.1.στ. Συχνότητες μεταλλάξεων νουκλεοτιδίων.
- III.1.ζ. Monte Carlo hypergeometric test για τη σύγκριση των φασμάτων μεταλλάξεων των νουκλεοτιδίων 752-800.
- III.1.η. Monte Carlo hypergeometric test για τη σύγκριση των φασμάτων μεταλλάξεων των νουκλεοτιδίων 1002-1176.

##### 2. Χρήση όλων των εγγραφών των γενετικών μεταλλάξεων (1424) και όλων των σωματικών (21587)

- III.2.α. Κατηγορίες γενετικών και σωματικών μεταλλάξεων στο γονίδιο *p53*.
- III.2.β. Συσχέτιση φύλου με ηλικία εμφάνισης καρκίνου στα γενετικά και στα σωματικά κύτταρα.
- III.2.γ. Σύγκριση ποσοστού των δύο φύλων στις γενετικές και σωματικές μεταλλάξεις.
- III.2.δ. Σύγκριση μέσου όρου ηλικίας εμφάνισης νεοπλασμάτων στις γενετικές και σωματικές μεταλλάξεις ανάλογα με την κατηγορία μετάλλαξης.
- III.2.ε. Καταγραφές σιωπηρών σωματικών μεταλλάξεων στην αλληλουχία του γονιδίου.
- III.2.στ. Μέσος όρος ηλικίας εμφάνισης καρκίνου στις γενετικές (α), στις σωματικές μεταλλάξεις (β) και στο γενικό πληθυσμό (γ).

- III.2.ζ. Συσχέτιση είδους καρκίνου και μεταλλάξεων που οφείλονται σε μεθυλίωση στα γενετικά και σωματικά κύτταρα.
- III.2.η. Συσχέτιση είδους καρκίνου με το είδος μετάλλαξης στις γενετικές και σωματικές μεταλλάξεις.
- III.2.θ. Μεταλλάξεις κωδικίων στην αλληλουχία του p53 ανάλογα με το είδος καρκίνου, στα γενετικά και σωματικά κύτταρα.
- III.2.ι. Καταγραφές σιωπηρών και υπόλοιπων μεταλλάξεων σε σχέση με το είδος του καρκίνου στα σωματικά κύτταρα.

### **3. Χρήση όλων των καταγραφών των σωματικών μεταλλάξεων (21587) και μία καταγραφή ανά οικογένεια των γενετικών μεταλλάξεων (282)**

- III.3.α. Τύποι γενετικών και σωματικών μεταλλάξεων.
- III.3.β. Η σταθμισμένη μεταλλαξιμότητα των γεννητικών κυττάρων με μοναδικές εγγραφές ανά οικογένεια
- III.3.γ. Κατανομή των μεταλλάξεων στην αλληλουχία του DNA στις γενετικές και σωματικές μεταλλάξεις.
- III.3.δ. Μεταλλαγμένα αμινοξέα σε γενετικό και σποραδικό καρκίνο.
- III.3.ε. Τύποι μεταλλάξεων των κωδικίων στην αλληλουχία του p53 στα εξώνια 5-8.
- III.3.στ. Απεικόνιση χωριστά των μεταλλάξεων κάθε νουκλεοτιδίου στην αλληλουχία του γονιδίου στα εξώνια 5-8.
- III.3.ζ. Σύγκριση του τύπου των μεταλλάξεων συνολικά για κάθε κωδικίο.
- III.3.η. Κατηγορίες των μεταλλάξεων συνολικά ανά κωδικίο.
- III.3.θ. Κατηγορίες μετάλλαξης των κωδικίων στην αλληλουχία του γονιδίου.
- III.3.ι. Συχνότητα και κατηγορία μετάλλαξης του κάθε αμινοξέως.
- III.3.κ. Κατηγορίες μεταλλάξεων σε σχέση με την αστάθεια που έχει αποδωθεί από τη μεθυλίωση της κυτοσίνης και της γουανίνης.
- III.3.κα. Τύποι μεταλλάξεων που πιθανώς σχετίζονται ή όχι με τη μεθυλίωση της κυτοσίνης.

### **4. Χρήση μόνο του συνόλου των γενετικών μεταλλάξεων (1424)**

- III.4.α. Συχνότητα μεταλλάξεων ανά γενεά.
- III.4.β. Κατηγορίες μεταλλάξεων ανά γενεά.
- III.4.γ. Μέσος όρος ηλικίας εμφάνισης καρκίνου ανά γενεά.
- III.4.δ. Συχνότητα μετάλλαξης κωδικίων στην αλληλουχία του γονιδίου ανά γενεά.
- III.4.ε. Συσχέτιση ορισμένων κωδικονίων με ηλικία εμφάνισης νεοπλασματος.
- III.4.στ. Είδη καρκίνου που εμφανίζονται ανά γενεά.
- III.4.ζ. Μέσος όρος ηλικίας εμφάνισης νεοπλασίας στα γενετικά κύτταρα.
- III.4.η. Συχνότητα ειδών καρκίνου ανάλογα με την φυλετική κληρονομία.
- III.4.θ. Ποσοστά μεταλλάξεων κάθε γενετικού συνδρόμου για κάθε είδος καρκίνου.
- III.4.ι. Ηλικία διάγνωσης του καρκίνου σε σχέση με τα γενετικά σύνδρομα.
- III.4.κ. Πλήθος καταχωρίσεων στη γενετική βάση για το κάθε είδος συνδρόμου.
- III.4.κα. Μεταλλάξεις κάθε συνδρόμου στην αλληλουχία του γονιδίου.

### **5. Χρήση μόνο της βάσης σωματικών μεταλλάξεων**

- III.5.α. Συχνότητες νεοπλασμάτων σε σχέση με τους καπνιστές και μη καπνιστές.
- III.5.β. Συχνότητες νεοπλασμάτων σε σχέση με την κατανάλωση οίνοπνεύματος.
- III.5.γ. Συνολικές καταγραφές στη βάση δεδομένων σε σχέση με τις επιδράσεις του περιβάλλοντος.

#### **IV. ΣΥΖΗΤΗΣΗ**

- IV.α. Ενδείξεις για το βαθμό μεθυλίωσης του γονιδίου.
- IV.β. Αξιολόγηση των προγραμμάτων ανάλυσης που χρησιμοποιήθηκαν στην εργασία.
- IV.γ. Σταθμισμένη μεταλλαξιμότητα.
- IV.δ. Γενετικός κώδικας.
- IV.ε. Η συμβολή των CpG θέσεων στη μεταλλαξιγένεση.
- IV.στ. Μέσος όρος ηλικίας εμφάνισης καρκίνου των σιωπηρών σωματικών μεταλλάξεων.
- IV.ζ. Οι γενετικές μεταλλάξεις των ενδονίων.
- IV.η. Πιθανοί μηχανισμοί που προάγουν ή αποτρέπουν τη μεταλλαξιγένεση.
- IV.θ. Διεισδυτικότητα των γενετικών μεταλλάξεων.
- IV.ι. Μεταλλάξεις δίχως οικογενειακό ιστορικό.
- IV.κ. Γενικά συμπεράσματα.

#### **V. ΒΙΒΛΙΟΓΡΑΦΙΑ**

##### **ΠΑΡΑΡΤΗΜΑ**

Ο Γενετικός Κώδικας  
Ξενόγλωσση Ορολογία

## I. ΕΙΣΑΓΩΓΗ

Με την ολοκλήρωση του Προγράμματος αποκρυπτογράφησης του Ανθρώπινου Γονιδιώματος έχει σχεδόν καταρτιστεί η λίστα των γονιδίων που χρειάζονται για να δημιουργηθεί ένας άνθρωπος. Όμως η δημιουργία είναι κατά πολύ πολυπλοκότερη από έναν απλό κατάλογο γονιδίων. Ίσης αξίας και σημασίας είναι ένα δεύτερο σύστημα που χρησιμοποιούν τα κύτταρα για να αποφασίσουν πότε και για πόσο χρόνο θα εκφράσουν τα γονίδια κατά την εξέλιξή τους. Επικρατεί σήμερα η άποψη πως η ρύθμιση αυτή επιτυγχάνεται μερικώς από τη μεθυλίωση της κυτοσίνης του CpG island των γονιδίων, που έχει ως αποτέλεσμα την απενεργοποίηση των γονιδίων.

Η μεθυλίωση του DNA στα σπονδυλωτά αποτελεί μία μέθοδο καταστολής της μεταγραφής με αποτέλεσμα να μεταγράφονται μόνο τα γονίδια που χρειάζονται τη συγκεκριμένη χρονική στιγμή. Αυτό δίνει την ευχέρεια στα σπονδυλωτά να συσσωρεύουν γονίδια που μεταγράφονται μόνο σε ορισμένα στάδια ανάπτυξης του οργανισμού ή χρειάζονται οι αντίστοιχες πρωτεΐνες σε μικρές ποσότητες και να τα χρησιμοποιούν εκλεκτικά δίχως να επιβαρύνουν τις λειτουργίες του κυττάρου συνέχεια. [1] [2]

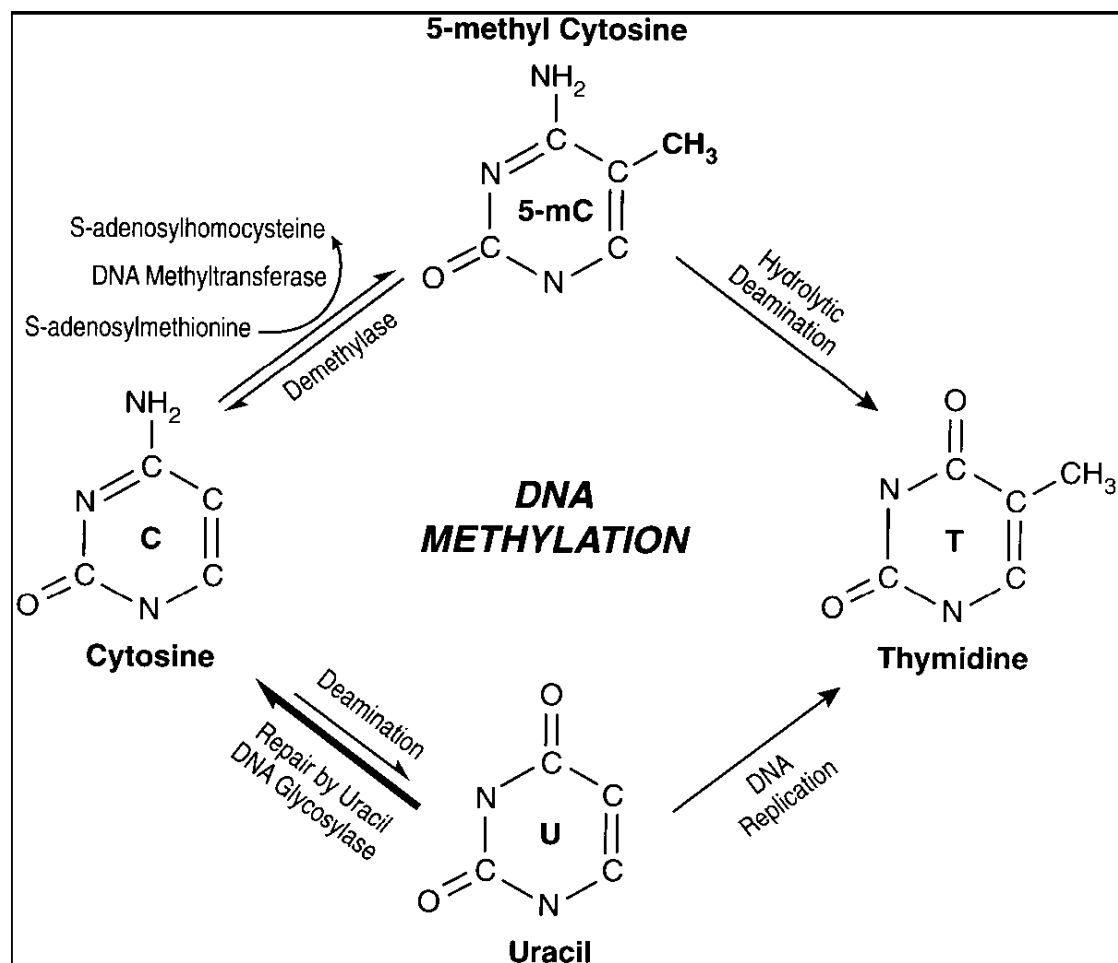
### Η μεθυλίωση της κυτοσίνης

Από την αναγνώρισή του το 1948, η Πέμπτη βάση του ανθρώπινου DNA 5-μεθυλο-κυτοσίνη (5-mC) έχει εγείρει το ενδιαφέρον των επιστημόνων και έχει εγκαινιάσει καινούργιο πεδίο επιστημονικών διχογνωμιών. Η μεθυλίωση του DNA στα ευκαρυωτικά κύτταρα σχετίζεται με την προσθήκη μιας μεθυλομάδας στον 5<sup>ο</sup> άνθρακα του δακτυλίου της κυτοσίνης. Η αντίδραση καταλύεται από την DNA μεθυλοτρανσφεράση με κατεύθυνση 5' -CG-3' και το δινουκλεοτίδιο που προκύπτει γράφεται σχηματικά ως CpG. Αποτελεί την πιο συχνή επιγενετική τροποποίηση στο ευκαρυωτικό DNA. Με τον όρο «επιγενετική» εννοούμε την τροποποίηση στην έκφραση των γονιδίων δίχως αλλαγή στην νουκλεοτιδική αλληλουχία. Οι επιγενετικές αλλαγές μπορεί να είναι κληρονομικές. [1] [3] [4] Η πιο γνωστή επιγενετική τροποποίηση είναι η μεθυλίωση της κυτοσίνης στη 5<sup>η</sup> θέση άνθρακα των CpG δινουκλεοτιδίων. [5]

Οι μη-μεθυλιωμένες κυτοσίνες μπορούν να υποστούν απαμίνωση σε ουρακίλη και στη συνέχεια η ουρακίλη-DNA γλυκοσιδάση επιδιορθώνει την αναντιστοιχία G:U αλλά όχι την G:T. Η DNA-μεθυλο-τρανσφεράση μπορεί να εμποδίσει την επιδιόρθωση και να συμβάλλει στην C>U>T μετάβαση. Η DNA-μεθυλοτρανσφεράση μπορεί ακόμη να μεσολαβήσει στην 5-mC>T transition στα CpG δινουκλεοτίδια όταν συνυπάρχουν συνθήκες αυξημένης έκφρασής



της ή όταν υπάρχει έλλειψη κυτταρικής S-αδενοσινο-μεθειονίνης. [6] [7]



**Εικόνα 1.** Σχηματική αναπαράσταση των βιοχημικών μονοπατιών της μεθυλίωσης της κυτοσίνης, της απομεθυλίωσης και της μεταλλαγμότητας της κυτοσίνης και της μεθυλοκυτοσίνης. [1]

### Γενικά για την CpG μεθυλίωση

Το δινουκλεοτίδιο CpG φαίνεται πολύ μικρό για να είναι βιολογικά σημαντικό, όμως οι ιδιότητες του είναι μοναδικές. Πάνω από 70% των βάσεων κυτοσίνης που συνδέονται με φωσφοδιεστερικό δεσμό με γουανίνη υπόκεινται σε 5' μεθυλίωση. Μ' αυτό τον τρόπο οι μεθυλιωμένες κυτοσίνες είναι απαμινώνονται σε θυμίνη, ενώ οι αμεθυλιωτές κυτοσίνες απαμινώνονται σε ουρακίλη, όπως φαίνεται στην εικόνα 1. [8] Περίπου 3% - 4% όλων των βάσεων κυτοσίνης στα θηλαστικά είναι μεθυλιωμένες και οι περισσότερες από τις 5-μεθυλιωμένες κυτοσίνες βρίσκονται στη δινουκλεοτιδική αλληλουχία CpG. [9] [10]

Υδρολυτική απαμίνωση συμβαίνει στις κυτοσίνες της διπλής έλικας του DNA με σχετικά αργό ρυθμό και ημιζωή περίπου 30.000 χρόνια

στους 37° και σε pH=7,4. Η μεθυλίωση στη 5<sup>η</sup> θέση της κυτοσίνης διευκολύνει την απαμίνωσή της. Η 5-μεθυλοκυτοσίνη στη διπλή έλικα του DNA απαμινώνεται 2,2 φορές γρηγορότερα σε σχέση με την κυτοσίνη. Φαίνεται ότι η διπλάσια ενίσχυση δεν είναι αρκετά γρήγορη για να εξηγήσει τα αυξημένα επίπεδα μεταλλάξεων των CpG δινουκλεοτιδίων, τα οποία υπολογίζονται 40 φορές γρηγορότερα από τις αμεθυλιώτες κυτοσίνες. [11] [12] [13] [14] [15] [16] [17]

Είναι πιθανό να υπάρχει διαφορετική επιδιόρθωση των αναντιστοιχιών T/G και U/G που προκύπτουν μετά την απαμίνωση. Η ουρακίλη φαίνεται ότι επιδιορθώνεται περισσότερο από τις ουρακίλο-DNA-γλυκοσιδάσες, ενώ η θυμίνη απομακρύνεται λιγότερο. [18] [19] [20] [21]

Όμως υπάρχει και η αντίθετη άποψη στην παραπάνω θεωρία, διότι ο ρυθμός των CpG μεταλλάξεων στις γενετικές σειρές των πρωτευόντων είναι 1250 φορές πιο αργός σε ένα τμήμα της αλληλουχίας του *p53* ιντρονίου 6, που εμφανίζει ομολογία με Alu αλληλουχίες και αποτελεί πιθανότατα μεταγενέστερη προσθήκη στο γονίδιο, σε σύγκριση με την απαμίνωση της 5-μεθυλοκυτοσίνης στη διπλή έλικα του DNA *in vitro*. Η μετάλλαξη της κυτταρικής σειράς από γενετικά κύτταρα είναι ακόμη πιο αργή στα CpG του παράγοντα IX. Αυτό δείχνει ότι *in vivo* επιδιορθωτικοί μηχανισμοί μπορούν να επιδιορθώνουν αποτελεσματικά όχι μόνο U/G αλλά και T/G. Παρόλα αυτά ελαττώματα στα ένζυμα της θυμινικής-DNA-γλυκοσιδάσης μπορεί να συμβάλλουν στην αύξηση των μεταλλάξεων των όγκων. [22] [23] [24]

Έχει υπολογιστεί ότι μόνο δύο 5-μεθυλοκυτοσίνες παθαίνουν απαμίνωση κάθε μέρα σε κάθε κύτταρο. Αυτό είναι σχετικά μικρό νούμερο, αν αναλογιστούμε ότι κάθε κύτταρο υπόκειται σε 100.000 προσβολές από οξειδωτικό στρες καθημερινά. Για να είναι σωστή η θεωρία της απαμίνωσης πρέπει να υποθέσουμε ότι άγνωστα ενδογενή μεταλλαξιγόνα προωθούν δραματικά την απαμίνωση των CpGs. [25] [26]

Το βενζο[α]πυρένιο είναι ισχυρό καρκινογόνο και βρίσκεται στα προϊόντα του καπνού σχηματίζοντας διμερή. Με μεταβολική ενεργοποίηση μετατρέπεται σε βενζο[α]πυρένιο διολεποξειδίο (BPDE) και προκαλεί G>T μεταλλάξεις. [27] [28] [29]

Πειράματα που διεξήχθησαν για τους μηχανισμούς επιδιόρθωσης του DNA ανάλογα με την αλυσίδα του DNA και ανάλογα με την αλληλουχία των νουκλεοτιδίων, έδειξαν ότι η μη-μεταγραφόμενη αλυσίδα επιδιορθώνεται 2-4 φορές λιγότερο αποτελεσματικά σε σχέση με αυτή που μεταγράφεται. Οι τοξίνες του βενζοπυρενίου

προσβάλλουν σε ίσο ποσοστό και τις δυο αλυσίδες. Η βλάβη αυτή απομακρύνεται πιο αργά στα CpG hotspots της μη-μεταγραφόμενης αλυσίδας σε σχέση με την μεταγραφόμενη. Τα ευρήματα δείχνουν ότι η αποτελεσματικότητα της επιδιόρθωσης παίζει καθοριστικό ρόλο στην πορεία της μεταλλαξιγένεσης. [30]

Επίσης μία εναλλακτική θεωρία προτείνει πως η μεθυλίωση της κυτοσίνης στα θηλαστικά είναι ένα σύστημα προστασίας του γενετικού υλικού από την εισβολή ξένης αλληλουχίας. Η μεθυλίωση της κυτοσίνης απενεργοποιεί τον προαγωγέα πολλών ιών και τρανσποζονίων. Σ' αυτά συμπεριλαμβάνονται οι ρετροϊοί και τα Alu στοιχεία. Η πλειονότητα των μεθυλιωμένων κυτοσινών βρίσκονται μέσα σε τέτοια στοιχεία. Απομεθυλιωτικά φάρμακα ενεργοποιούν τη μεταγραφή ενδογενών τρανσποζονίων. [1] [31]

Έχει παρατηρηθεί και σε κυτταροκαλλιέργειες που έχουν διαμολυνθεί με ξένο DNA και σε οργανισμούς στους οποίους γίνεται εισαγωγή ξένου DNA ότι στην αρχή το καινούργιο DNA που ενσωματώνεται υφίσταται μεθυλίωση. Άρα η καινούργια μεθυλίωση συνιστά ένα φράγμα αμυντικό για να κατασταλεί η έκφραση του ξένου DNA. Χαρακτηριστικό παράδειγμα αποτελεί η de novo μεθυλίωση του ιού του έρπητα στα θηλαστικά μετά την ενσωμάτωσή του στο γονιδίωμα του ξενιστή. [32] [33]

Η μετατροπή από κυτοσίνη σε μεθυλοκυτοσίνη προάγει το σχηματισμό διμερών διπυριμιδίνης κατά 15 φορές όταν υπάρχει έκθεση στο ηλιακό φως. [34]

### **Οι μεθυλοτρανσφεράσες**

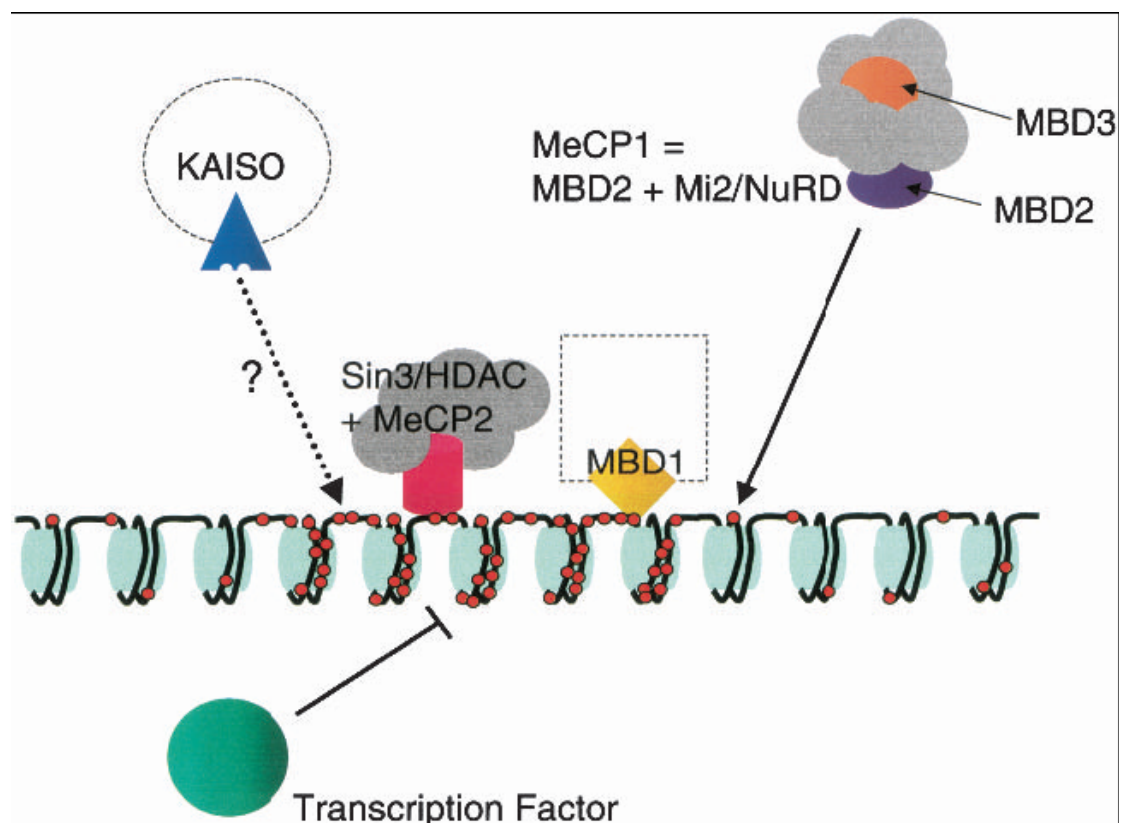
Η μηχανή μεθυλίωσης των θηλαστικών αποτελείται από δύο στοιχεία: τις DNA-μεθυλοτρανσφεράσες (DNMTs), οι οποίες είναι υπεύθυνες για την εγκατάσταση και τη διατήρηση των προτύπων μεθυλίωσης, και τις μεθυλ-CpG-συνδεόμενες πρωτεΐνες (MBDs), οι οποίες σχετίζονται με το «διάβασμα των σημαδιών» της μεθυλίωσης. Η μεθυλίωση του DNA καταστέλλει τη μεταγραφή άμεσα, αποτρέποντας τη σύνδεση συγκεκριμένων μεταγραφικών παραγόντων, και έμμεσα, προσελκύοντας τις μεθυλ-CpG-συνδεόμενες πρωτεΐνες. Οι μεθυλ-CpG-συνδεόμενες πρωτεΐνες ξεκινούν τον καταρράκτη της συμπυκνωμένης επαναδιαμόρφωσης της χρωματίνης. [35] [36]

Υπάρχουν τρεις κατηγορίες DNA μεθυλοτρανσφερασών: DNMT1, DNMT2, DNMT3(a και b). [37] [38] Τα υπολείμματα κυτοσίνης που είναι ασυμμετρικά μεθυλιωμένα αποτελούν το υπόστρωμα της DNMT1, η οποία είναι υπεύθυνη για τη διατήρηση της συμμετρικής μεθυλίωσης του γονιδιώματος. [39] Δεν έχει διευκρινιστεί επαρκώς ο

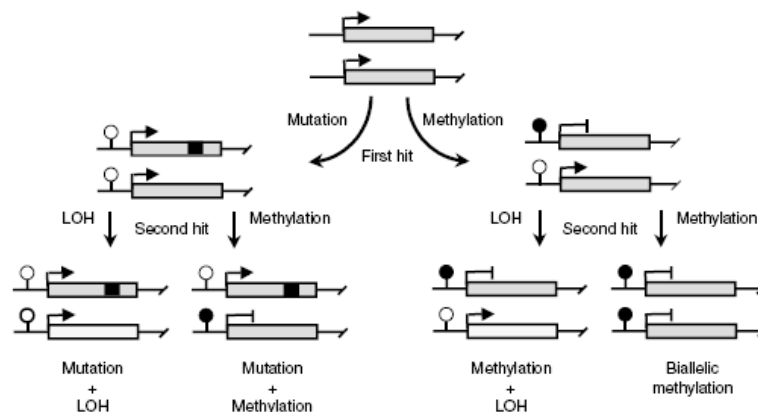
ρόλος της DNMT2, ενώ γνωρίζουμε ότι η DNMT3 (a και b) είναι υπεύθυνη για την de novo μεθυλίωση των υπολειμμάτων CpG, CpA, CpT, CpC τόσο στα εμβρυικά βλαστικά κύτταρα όσο και στα σωματικά κύτταρα. [40] [41]

Λίγες είναι οι πληροφορίες για τον τρόπο που η μεθυλίωση του DNA πραγματοποιείται στοχευμένα σε συγκεκριμένες γενετικές θέσεις. Πιθανότατα αυτό γίνεται χάρη στη διάδραση μεταξύ DNA-μεθυλοτρανσφερασών και των πρωτεϊνών της χρωματίνης. Μελέτες με knockout ποντίκια έδειξαν ότι οι DNMT1 DNMT3b είναι απαραίτητες για την εμβρυική ανάπτυξη, ενώ ποντίκια με έλλειψη της DNMT3a πεθαίνουν μερικές εβδομάδες μετά τη γέννηση. [42] [43]

Ο προαγωγέας του γονιδίου είναι πολύ σημαντικός για τη διατήρηση της αμεθυλιώτης νησίδας CpG. Τα γονίδια που μεταφέρονται σε ζώα και φέρουν CpG νησίδες διπλασιάζονται μαζί με τα κύτταρα και εκφράζονται διαρκώς, εκτός αν χάσει ο προαγωγέας τη λειτουργικότητά του. Τότε τα ξένα γονίδια μεθυλιώνονται και στα CpG νησίδια και σταματά η έκφρασή τους. [44] [45] Παρομοίως ιικό DNA που ενσωματώνεται στα εμβρυικά βλαστικά κύτταρα μεθυλιώνεται όταν απουσιάζει ενισχυτής από τον προαγωγέα, ενώ αποβάλλει τη μεθυλίωση όταν ο ενισχυτής είναι παρών. [46]



**Εικόνα 2.** Μηχανισμοί της μεταγραφικής καταστολής λόγω της μεθυλίωσης του DNA. Η αλυσίδα του DNA απεικονίζεται με μαύρο χρώμα και οι μεθυλιωμένες κυτοσίνες με κόκκινο. Κάτω από την αλυσίδα απεικονίζεται ένας μεταγραφικός παράγοντας με πράσινο κύκλο, ο οποίος δεν μπορεί να συνδεθεί στην περιοχή αναγνώρισής του αν υπάρχει μεθυλιωμένο δινοκυλεοτίδιο CpG μέσα σ' αυτή. Πάνω από το DNA βρίσκονται τα σύμπλοκα που έλκονται από τη μεθυλίωση. Αυτά είναι το MeCP1, το MeCP2, το MBD1 και το KAISO. Ακόμη οι μηχανισμοί των αλληλεπιδράσεων δεν είναι πλήρως κατανοητοί. [5]



**Εικόνα 3.** Η υπόθεση δύο χτυπημάτων του Knudson όπως επαναδιατυπώθηκε από τους Jones and Laird. Στην αρχική υπόθεση, τα δύο ενεργά αλληλία ενός ογκοκατασταλτικού γονιδίου απενεργοποιούνται από μεταλλάξεις και απώλεια της ετεροζυγωτίας αντίστοιχα. Η μεταγραφική καταστολή εξαιτίας μεθυλίωσης του προαγωγέα του γονιδίου μπορεί να συμβεί σε οποιοδήποτε στάδιο και αντιπροσωπεύει τρίτο μονοπάτι γονιδιακής απενεργοποίησης.

Αυτό που οι παλαιότεροι ερευνητές ονόμαζαν Loss of Heterozygosity πίστευαν ότι οφείλεται σε απώλεια χρωμοσωμικού τμήματος. Είναι όμως εξίσου πιθανό να οφείλεται σε μεθυλίωση του CpG island και απενεργοποίηση του γονιδίου. [8]

### Το γονίδιο *p53*

Κατά τη διάρκεια της ζωής τους τα κύτταρα εκτίθενται διαρκώς σε ενδογενή και εξωγενή στρες, τα οποία επηρεάζουν τη συμπεριφορά των κυττάρων. Προσβολές του γενετικού υλικού των κυττάρων είναι ιδιαίτερα βλαπτικές διότι μπορούν να μεταδοθούν στα θυγατρικά κύτταρα και να οδηγήσουν σε νεοπλασίες. Για να εξασφαλιστεί η ομοιόσταση, τα κύτταρα των θηλαστικών έχουν ρυθμιστικές πρωτεΐνες-κλειδιά που ελέγχουν την ομαλή κυτταρική ανάπτυξη. Αρχικά η πρωτεΐνη *p53* θεωρήθηκε ως ογκοκατασταλτική διότι είναι απαραίτητη για τον έλεγχο της βλάβης του DNA. Αν η βλάβη είναι περιορισμένη τότε σταματά η αντιγραφή και μεταγραφή του

γενετικού υλικού μέχρι να επιδιορθωθεί. Αν η βλάβη είναι μεγάλη τότε ενεργοποιεί μονοπάτια απόπτωσης. Η πρωτεΐνη *p53* είναι λειτουργική ως τετραμερές και βρίσκεται σε πολύ μικρά ποσά σε φυσιολογικά κύτταρα δίχως στρες. Αν όμως το κύτταρο υποστεί στρες, για παράδειγμα ενεργοποίηση ογκογονιδίων ή υποξία, τότε ενεργοποιούνται μονοπάτια που τροποποιούν μετα-μεταγραφικά και σταθεροποιούν τη δομή της πρωτεΐνης *p53*. [47]

Το γονίδιο *p53* θεωρείται ως το πιο συχνά μεταλλάσσόμενο γονίδιο στους ανθρώπινους καρκίνους και το φάσμα των μεταλλάξεων του έχει αναλυθεί σε σχέση με πολλούς παράγοντες. [48] [49] [50] [51] Το γονίδιο *p53* μεταλλάσσεται σε ποσοστό άνω του 50% στο σύνολο των ανθρώπινων καρκίνων. Αυτό δείχνει τον σημαντικότερο ρόλο που διαδραματίζει στον έλεγχο της καλής λειτουργίας του κυττάρου και της ακεραιότητας του γονιδιώματος. Η μεταλλαγμένη πρωτεΐνη συνήθως χάνει την ικανότητα να συνδέεται με το DNA ή να σχηματίζει διμερή και τετραμερή. [52]

Είναι αξιοσημείωτο, ότι μόνο η έλλειψη του *p53* σχετίζεται με την εμφάνιση νεοπλασμάτων. Πολύ λίγες μεταλλάξεις στα ομόλογα γονίδια *p63*, *p73* έχουν ανιχνευθεί στα καρκινικά κύτταρα. Επίσης τα *p63*, *p73* knockout ποντίκια έχουν αναπτυξιακά ελαττώματα, αλλά δεν παρουσιάζουν αύξηση της συχνότητας όγκων. [53]

Οι πρωτεΐνες της οικογένειας του *p53* έχουν την μεγαλύτερη διατήρηση εξελικτικά στην κεντρική περιοχή που δεσμεύει το DNA, ενδιάμεσα στο C-τελικό άκρο του τετραμερισμού και τη μικρότερη στην N-τελική περιοχή. [54] [55] [56] [57] [58] Η ανάλυση του φάσματος των μεταλλάξεων αποκαλύπτει χαρακτηριστικές θέσεις αυξημένης μεταλλαξιμότητας που ονομάζονται hotspots. Αυτά διαφέρουν ανάλογα με το είδος του καρκίνου και επηρεάζονται από εξωγενή καρκινογόνα. [59] [60]

Το γονίδιο *p53* θεωρείται πλήρως μεθυλιωμένο σε θέσεις CpG. Ταυτόχρονα η πλειοψηφία των μεταλλάξεων εντοπίζονται στα δινουκλεοτίδια CpG δημιουργώντας σημεία αλληλουχίας υψηλής μεταλλαξιμότητας (mutation hotspots). Η γενετική αστάθεια θεωρείται ότι προέρχεται από την απαμίνωση της μεθυλιωμένης κυτοσίνης με άμεσο σχηματισμό μιας βάσης θυμίνης. Όμως η συχνότητα των CpG μεταλλάξεων ποικίλλει στα διάφορα είδη καρκίνου. Άρα και άλλα στοιχεία, εκτός από την παρουσία της μεθυλίωσης επηρεάζουν το προφίλ των μεταλλάξεων σε θέσεις CpG. [51] [61] Τα περισσότερα hotspots εντοπίζονται σε θέσεις CpG και αφορούν C>T και G>A μεταβάσεις και G>T μεταστροφές. Αυτές έχουν συσχετισθεί με την μεθυλίωση της κυτοσίνης. [51] [62]

Απενεργοποίηση του *p53* γονιδίου προκαλείται εξαιτίας μικρής έκτασης μεταλλάξεων του γονιδίου (missense, nonsense, εισαγωγή ή διαγραφή λίγων νουκλεοτιδίων). 90% παρατηρείται παραγωγή μεταλλαγμένης πρωτεΐνης *p53* και 10% παρατηρείται απουσία πρωτεΐνης *p53*. Επίσης είναι δυνατό να απενεργοποιηθεί εξαιτίας υπερμεθυλίωσης του CpG νησιδίου του προαγωγέα. [47]

### **Τι πρότυπα νουκλεοτιδίων μελετούμε και για ποιο λόγο**

Σημαντική γενετική αστάθεια σχετίζεται με συγκεκριμένα στοιχεία της αλληλουχίας του DNA. Αυτά τα στοιχεία είναι τα δι- και τρι-νουκλεοτίδια (CC,GG, CCC, GGG) τα οποία μπορούν να υποστούν επιγενετικές ή γενοτοξικές μετατροπές, και τα δινουκλεοτίδια GT/TG και CA/AC, τα οποία συμβάλουν στην αποσταθεροποίηση της δίκλωνης έλικας του DNA. Γενετική αστάθεια παρατηρείται ιδιαίτερα όταν συνδυάζονται μεταξύ τους. [63]

Οι αλληλουχίες του DNA: GG, GGG, CC, CCC καθώς και τα δινουκλεοτίδια TG, GT, CA, AC και οι επαναλήψεις τριπλέτας νουκλεοτιδίων έχουν παρουσιαστεί ως υπεύθυνοι για την προσωρινή μετατροπή της δίκλωνης έλικας του DNA σε εναλλακτικές ελικοειδείς διαμορφώσεις. Αυτό έχει ως αποτέλεσμα τη γενετική αστάθεια και την τροποποίηση της γενετικής πληροφορίας. Είναι πιθανό στην κληρονομική προδιάθεση για απενεργοποίηση του *p53* να αναπτύσσονται μεταλλάξεις κοντά σε ευάλωτες θέσεις CpG και GT, TG, CA, AC ή AT-πλούσιες περιοχές, οι οποίες προκαλούν αποσταθεροποίηση της δομής του DNA. Ή ακόμη και όταν κωδικοποιούν για αμινοξέα που ελάχιστα συμβάλλουν στη δραστηριότητα της *p53*, να συνδέονται με γενετική αστάθεια όταν το γονίδιο υφίσταται επιγενετική τροποποίηση. Ενώ στους σποραδικούς καρκίνους η απενεργοποίηση είναι πιθανό να οφείλεται κυρίως σε επιγενετικές και γενοτοξικές επιδράσεις στα πρότυπα CC, GG, CCC και GGG νουκλεοτιδίων. Στη γενετική αστάθεια βοηθά η ασύμμετρη μεθυλίωση των νουκλεοτιδίων και η αποσταθεροποίηση της δίκλωνης έλικας του DNA. [63]

Φαίνεται ότι προκαλούν κυρίως CG>AT μεταβάσεων ακόμη και σε μη μεθυλιωμένες αλληλουχίες CpG. Πρόσφατες μελέτες έδειξαν ότι μη-CpG μεθυλίωση παρατηρείται στον προαγωγέα του *p53* στις αλληλουχίες CCWGG και στα εξώνια στις αλληλουχίες CC και CCC. Οι αλληλουχίες αυτές είναι συμπληρωματικές με τις αλληλουχίες GG και GGG που σχηματίζουν O-6-μεθυλο-γουανίνη με την επίδραση γενοτοξικών παραγόντων. [64] [65]

Η συχνότητα των G:C>A:T μεταβάσεων σε αλληλουχίες non-CpG είναι υψηλή σε καρκίνους του μαστού και της κύστης αλλά διαφέρει σημαντικά ανάμεσα σε καρκίνους που σχετίζονται με το κάπνισμα.



Δηλαδή σε καρκίνους των βρόγχων ή της κεφαλής και τραχήλου παρατηρείται σχηματισμός διμερών της γουανίνης. Αυτό υποδηλώνει ότι πέρα από τις μεταβολές των προτύπων CpG, CC, GG, CCC, GGG, τα προϊόντα του καπνού και τα μεταλλικά ιόντα μπορούν να συμβάλλουν στη γενετική αστάθεια με την πρόκληση αλλαγών στη διαμόρφωση των μεθυλιωμένων αλληλουχιών. [66] Ωστόσο, η υψηλή συχνότητα των C>T μεταβάσεων που παρατηρείται σε non-CpG αλληλουχίες, ιδιαίτερα σε καρκίνους που σχετίζονται με το κάπνισμα, αποτελεί πεδίο δίχως ικανοποιητική μέχρι σήμερα απάντηση. [67]

Έχει πρόσφατα ανακοινωθεί η παρουσία non-CpG μεθυλίωσης σε αλληλουχίες CC, CCC για το *p53* εξώνιο 5 σε ασθενείς με καρκίνο του πνεύμονα και σε πρότυπα CCWGG. [65] [68] Οι CC, CCC αλληλουχίες είναι συμπληρωματικές των GG, GGG και οι τελευταίες θεωρούνται στόχος για το σχηματισμό διμερών O-6-μεθυλογουανίνης. [69] Οι Kouidou et al. έδειξαν ότι στις σιωπηρές μεταλλάξεις του *p53* στα εξώνια 5-8 η γενετική αστάθεια δεν είναι τυχαία, αλλά σχετίζεται με τις αλληλουχίες CC, CCC, GG, GGG και σε μικρότερο βαθμό με τις CpG που βρίσκονται δίπλα σε TG, GT, AC, CA δινουκλεοτίδια και στις επαναλήψεις τους. [70]

Εφόσον οι σιωπηρές μεταλλάξεις δεν επηρεάζουν την προκύπτουσα πρωτεΐνη, η συχνότητα των σιωπηρών μεταλλάξεων σχετάζεται μόνο με την ταχύτητα μεταλλαξιγένεσης του προτύπου νουκλεοτιδίων. Από το φάσμα των σιωπηρών σωματικών μεταλλάξεων του *p53* (σύνολο 232) αναλύθηκαν οι μεταλλάξεις των G:C>A:T μεταβάσεων και προκύπτει το συμπέρασμα ότι οι περισσότερες μεταλλάξεις οδηγούν στο σχηματισμό γονιδιακών προτύπων ευάλωτων σε μεθυλίωση CpG, CCC, GGG (65/232). Από τις υπόλοιπες, ένα σημαντικό ποσοστό σχηματίζει μικρές επαναλήψεις δι- και τρι-νουκλεοτιδίων ή παλίνδρομες αλληλουχίες (74/232). Οι μεταλλάξεις συμβαίνουν συνήθως σε θέσεις πιθανής μεθυλίωσης ή βρίσκονται κοντά σε αυτές. [71] Τα παραπάνω δείχνουν ότι οι non-G:C>A:T μεταλλάξεις σχετίζονται με την αύξηση των θέσεων πιθανής μεθυλίωσης, με το σχηματισμό διμερών της γουανίνης GGG, ή με την αύξηση της συμμετρίας της αλληλουχίας. Η αυξημένη ευαισθησία των CC αλληλουχιών σε μεταλλάξεις μπορεί να συσχετισθεί με την non-CpG μεθυλίωση της κυτοσίνης. [68] [72]

### **Τα δινουκλεοτίδια TG και CA αποτελούν θέσεις αυξημένης γενετικής αστάθειας**

Επαναλαμβανόμενα στοιχεία στα εξώνια 5-8 του *p53* που είναι διακοπτόμενες αλληλουχίες οκτώ GT, GA, AC σχετίζονται περισσότερο με nonsense μεταλλάξεις. Αυτά τα δινουκλεοτίδια προκαλούν γενετική αστάθεια και σε άλλους οργανισμούς. [73] Με



την ανάλυση των μεταλλάξεων του *p53* φαίνεται ότι και μικρότερες επαναλήψεις ή παλίνδρομες αλληλουχίες μπορούν να επιφέρουν γενετική αστάθεια. [71]

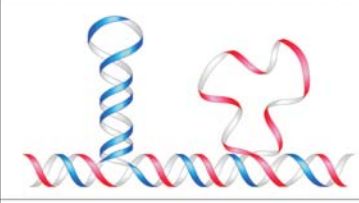

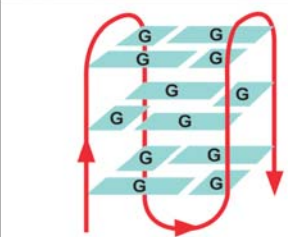
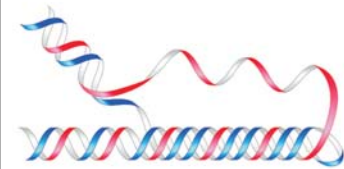
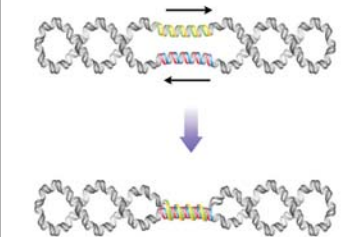
Η *Drosophila Melanogaster* και ορισμένα είδη ψαριών έχουν κοινά χαρακτηριστικά στην αλληλουχία του DNA παρόλο που ανήκουν σε διαφορετική κατηγορία στην ταξινόμηση των οργανισμών. Τα δινουκλεοτίδια TG και CA παρεμβάλλονται μεταξύ AT-πλούσιες περιοχές και δρουν ως στόχος για τις High-Mobility Group (HMG) πρωτεΐνες. Τα δινουκλεοτίδια TG και CA έχουν δομική αμφισταθερότητα και έχουν ευχέρεια για πολλές μορφές ελικώσεων. Επομένως τα TG και CA δινουκλεοτίδια με τη βοήθεια των παραπάνω πρωτεϊνών μπορούν να προκαλέσουν ιδιαίτερη αναδίπλωση της έλικας του DNA. [74] [75]

Οι παλινδρομικές αλληλουχίες επιτρέπουν σημαντικές μεταβολές της διαμόρφωσης του DNA. Αυτές μπορεί να προκαλέσουν δομική μετατροπή του DNA σε τρίκλωνη φουρκέτα, ιδιαίτερα σε περιοχές που συσσωρεύουν μεγάλη αρνητική υπερελίκωση. Ο λειτουργικός ρόλος της τρίκλωνης φουρκέτας δεν είναι ακόμη γνωστός αλλά θεωρείται ότι μεσολαβεί ρυθμιστικά για την απορρόφηση της αρνητικής υπερελίκωσης. Θεωρείται ακόμη ότι οι ευλύγιστες αλληλουχίες βοηθούν τον ανασυνδυασμό του DNA κατά τη φάση του πολλαπλασιασμού του. [76] [77] [78] [79]

### **Ιδιότητες των τρινουκλεοτιδίων για σχηματισμό non-B DNA δομών**

Οι ιδιότητες των επαναλαμβανομένων τρινουκλεοτιδίων, τετρανουκλεοτιδίων και πεντανουκλεοτιδίων σχετίζονται με το σχηματισμό non-B DNA δομών που περιλαμβάνουν φουρκέτες (hairpins), ολίσθηση της έλικας (slipped strand), στοιχεία αποελίκωσης του DNA (unwinding elements), τρικλωνη έλικα (triplex), τετράκλωνη έλικα (tetraplex) και κολλώδες DNA (sticky DNA). Αυτές οι δομές θεωρείται ότι προκαλούν μερικές κληρονομικές νευρολογικές ασθένειες. [80] [81] [82]

Αλληλουχίες με τέτοιες ιδιότητες είναι CTG•CAG, CGG•CCG, GAA•TTC, GAC•GTC, CCTG•CAGG, ATTCT•AGAAT. Non-B DNA δομές σχηματίζονται κατά τον διπλασιασμό, την μεταγραφή, τον ανασυνδυασμό και την επιδιόρθωση του DNA. Επίσης non-B DNA δομές σχηματίζονται ως αποτέλεσμα της αρνητικής υπερελίκωσης του DNA. [83] [84] [85] [86] [87] [88] Αυτές οι δευτερογενείς δομές φαίνεται ότι είναι στόχοι για ορισμένα επιδιορθωτικά ένζυμα. [89] [90] [91] [92] [93] [94] [95]

Structure	Conformation	General Seq. Requirements	Sequence
Slipped (Hairpin) Structure		Direct Repeats	<pre> CNGCNGCNG GNCGNCGNC </pre>
DNA Unwinding Element		A-T rich Regions	<pre> ATTCTATTCT TAAGATAAGA </pre>
Tetraplex		Single-Strand Oligo (G) <sub>n</sub> Tracts	<pre> CGGCGGCGG GCCGCCGCC </pre>
Triplex		(R•Y) <sub>n</sub> Mirror Repeats	<pre> GAAGA AGAAG CTTCT TCTTC </pre>
Sticky DNA		2 G-A rich Tracts Direct Repeats	<pre> GAAGAAGAA CTTCTTCTT </pre>

**Εικόνα 4.** Παρατηρούμε τις εναλλακτικές διαμορφώσεις που δημιουργούν τα διάφορα πρότυπα νουκλεοτιδίων. Οι αλληλουχίες με δομή (CNG)<sup>n</sup> και γενικά αλληλουχίες με επαναλαμβανόμενα δι- και τρινουκλεοτίδια σχηματίζουν φουρκέτες και σταυρούς. Οι αλληλουχίες που είναι πλούσιες σε A και T είναι ευάλωτες στην αποελίκωση του DNA. Οι αλληλουχίες που έχουν συμμετρία καθρέφτη σχηματίζουν δομές τρίκλωνης έλικας. Οι αλληλουχίες που είναι πλούσιες σε G κυρίως στη μια αλυσίδα είναι επιρρεπείς στο σχηματισμό τετράκλωνης διαμόρφωσης. Οι αλληλουχίες που είναι πλούσιες σε G και A κυρίως σε μία αλυσίδα και περιάχουν και επαναλαμβανόμενα πρότυπα νουκλεοτιδίων μπορούν να σχηματίσουν κολλώδες DNA. [96]

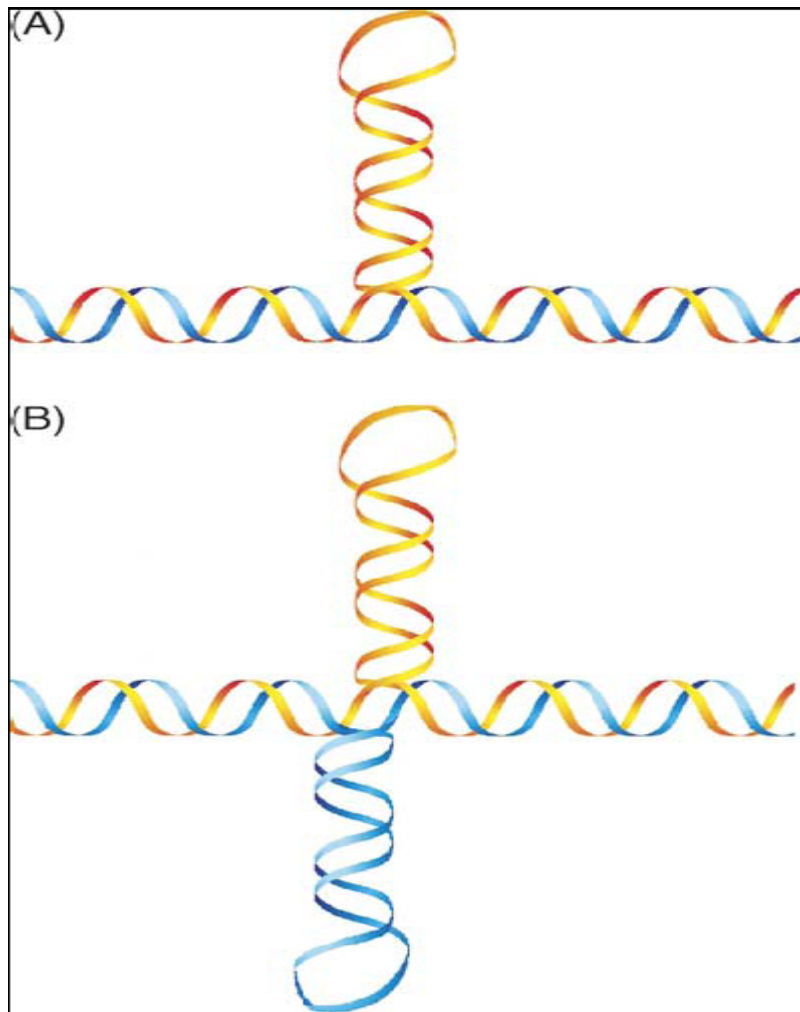
## Non B – DNA διαμορφώσεις

Σήμερα έχουν βρεθεί τουλάχιστον 10 διαφορετικά είδη από non-B DNA διαμορφώσεις και είναι πιθανό να ανακαλυφθούν και άλλες non-B DNA διαμορφώσεις στο μέλλον. Αντίστροφες επαναλήψεις μπορούν να σχηματίσουν δομές φουρκέτας. Ευθείες επαναλήψεις μπορούν να υιοθετήσουν ολισθαίνουσες δομές με εξέχουσες βάσεις. Επαναλαμβανόμενα στοιχεία πλούσια σε κατάλοιπα γουανίνης μπορούν να σχηματίσουν τετράκλωνη έλικα ολιγο-γουανίνης. Εναλλασσόμενες αλληλουχίες πυριμιδίνης και πουρίνης μπορούν να μεταπέσουν σε αριστερόστροφες Z-DNA δομές. Στοιχεία πολypουρίνης-πολυπυριμιδίνης με επαναλήψεις καθρέπτη μπορούν να υιοθετήσουν τρίκλωνη έλικα DNA (H-DNA). Αλληλουχίες που περιέχουν δύο μεγάλες επαναλήψεις GAA μπορούν να σχηματίσουν κολλώδες DNA, μία δομή από δύο τριπλές έλικες. Επίσης, A-T πλούσιες περιοχές του DNA, συνήθως συναντώνται στην αρχική περιοχή της μεταγραφής, υπό την πίεση του υπερελικωμένου DNA οδηγούν στο ξετύλιγμα της διπλής αλύσου. [97]

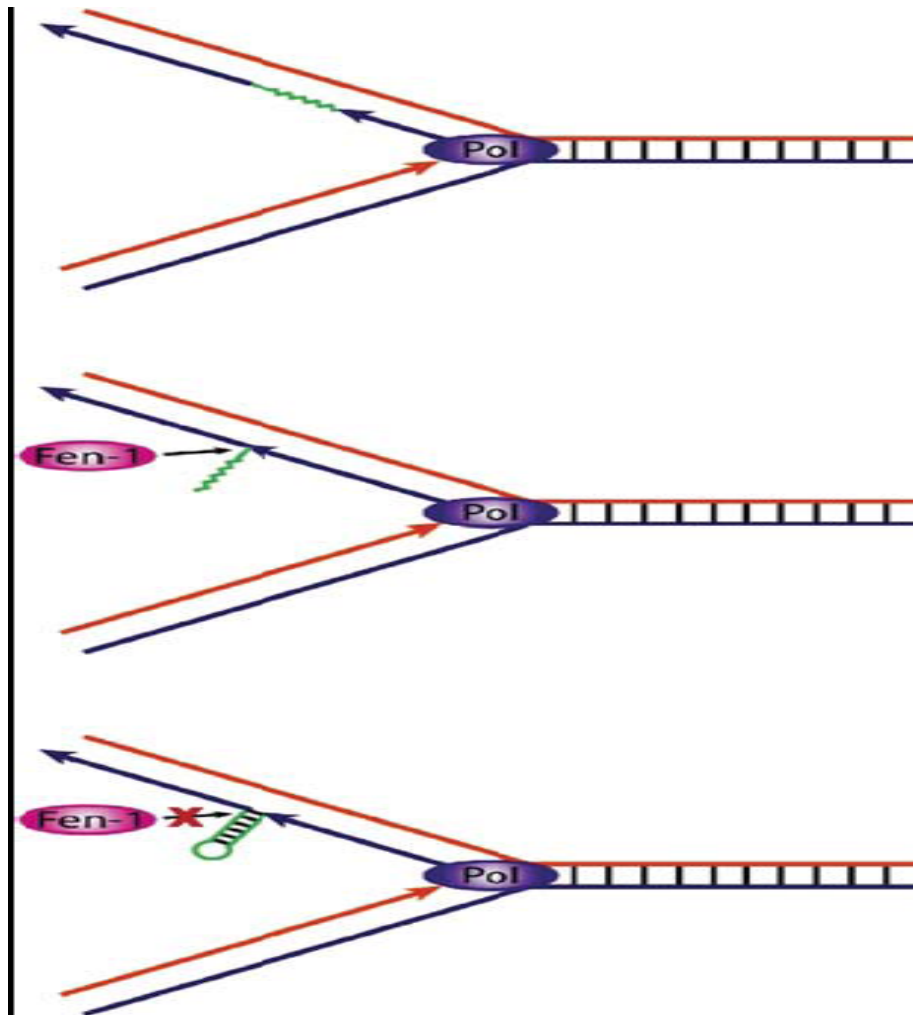
Στο ανθρώπινο γονιδίωμα, οι επαναλαμβανόμενες αλληλουχίες αποτελούν πάνω από το 50% του συνολικού γενετικού υλικού, όπου οι επαναλήψεις απλών αλληλουχιών (επαναλήψεις προτύπων από 1-6 βάσεις ονομάζεται μικροδορυφορικό DNA, επαναλήψεις μεγαλύτερων προτύπων από 7-100 βάσεις ονομάζονται μινιδορυφορικό DNA) συνθέτουν περίπου το 3% του συνολικού DNA. [98]

### Διαμόρφωση φουρκέτας και σταυρού

Η αντιγραφή του DNA, η μεταγραφή και η διαδικασία επιδιόρθωσης οδηγούν στο ξετύλιγμα της δίκλωνης έλικας, δίνοντας χρόνο στη μονή έλικα του DNA να αναδιπλωθεί γύρω από τον εαυτό της και να πάρει non-B DNA δομές. Οι αντίστροφες επαναλήψεις, αλλά και οι ευθείες επαναλήψεις, όπως του CNG τρινουκλεοτιδίου, σχηματίζουν φουρκέτες και σταυρούς. Στις αντίστροφες επαναλήψεις, 14 βάσεις συνολικά (7 ανά κατεύθυνση), είναι αρκετές για το σχηματισμό σταθερής φουρκέτας. Στους σχηματισμούς του CNG τρινουκλεοτιδίου κάθε δύο ζεύγη C-G παρεμβάλλεται ένα αναντίστοιχο ζεύγος. [99] [100]



**Εικόνα 5.** Δομές φουρκέτας και σταυρού. (A) Η δομή φουρκέτας σχηματίζεται στη μία έλικα του DNA στη θέση αντίστροφης επανάληψης προτύπου. (B) Η δομή του σταυρού αποτελείται από δύο φουρκέτες, μία φουρκέτα για κάθε αλυσίδα στην ίδια γενετική θέση. [99]



**Εικόνα 6.** Αυτό είναι ένα υποθετικό μοντέλο για την προσθήκη βάσεων στην έλικα του DNA. Η δομή φουρκέτας σχηματίζεται σε ένα τμήμα Okasaki. Η ενδονουκλεάση FEN1 μπορεί να κόψει την ασταθή φουρκέτα, όχι όμως και τη σταθερή. Αυτό έχει ως αποτέλεσμα την προσθήκη του τμήματος της φουρκέτας στο τελικό μεταγράφημα. [99] Οι δομές φουρκέτας και σταυρού είναι γνωστά υποστρώματα για ορισμένες ενδονουκλεάσες και για ένζυμα επιδιόρθωσης. Η δράση αυτών των ενζύμων προκαλεί σπάσιμο στη δίκλωνη έλικα του DNA και κατ' επέκταση τοπική γενετική αστάθεια. Μπορεί να προκύψει αναδιάταξη ή μετάθεση χρωμοσωμική, γεγονός που επιβεβαιώθηκε πειραματικά. [101] [102] [103]

## Z – DNA

Z-DNA είναι αριστερόστροφη έλικα με διαμόρφωση ζικ-ζακ στο σκελετό (backbone). Επειδή οι βάσεις πουρίνης μπορούν να πάρουν συνηθισμένη διαμόρφωση, ειδικές αλληλουχίες που έχουν εναλλασσόμενες περιοχές πουρίνης-πυριμιδίνης, όπως είναι GC και GT επαναλήψεις, μπορούν να πάρουν άμεσα Z-DNA δομή και *in vitro* και *in vivo*. [104] [105] Οι αλληλουχίες που μπορούν να πάρουν Z-διαμόρφωση είναι άφθονες στο ευκαρυωτικό γονιδίωμα και συναντώνται περίπου μία φορά κάθε 3000 βάσεις στο ανθρώπινο γονιδίωμα. Επίσης Z διαμόρφωση μπορούν να πάρουν οι μεθυλιωμένες αλληλουχίες. [106] [107]

Οι επαναλήψεις GT είναι τα πιο συχνά και απλά επαναλαμβανόμενα στοιχεία του ανθρώπινου γονιδιώματος. Σύμφωνα με εκτιμήσεις αυτές οι επαναλήψεις ξεπερνούν το 0,25% του συνολικού γονιδιώματος. [98]



**Εικόνα 7.** Η αριστερόστροφη ζικ-ζαγκ φύση του Z – DNA σκελετού εντοπίζεται στο κέντρο του μορίου, με κανονικά B – DNA άκρα. Μονές αλυσίδες 3-4 βάσεων σχηματίζονται στις συνδέσεις μεταξύ B – Z διαμόρφωσης. [99]

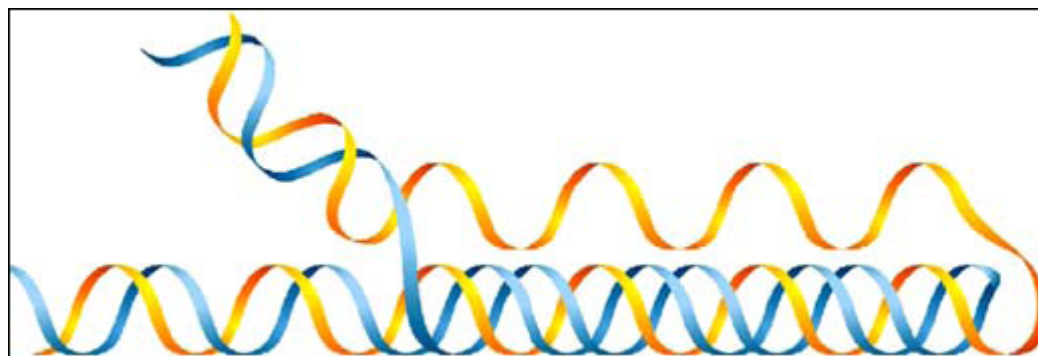
Οι παρεμβολές Z – μορφής σε μία έλικα DNA σχετίζονται με αύξηση της μεταλλαξιγένεσης. Έτσι, όταν τα πλασμίδια που περιέχουν CG σε ικανή ποσότητα (6-29 επαναλήψεις) για να σχηματίσουν Z-DNA πολλαπλασιάζονται σε βακτήρια, διαγραφές νουκλεοτιδίων συμβαίνουν με συχνότητα πολύ μεγαλύτερη από τα πλασμίδια που δεν περιέχουν τις CG επαναλήψεις. [108] [109] Αυτό το φαινόμενο φαίνεται να είναι ειδικό για τη δομή του Z-DNA και όχι για τη διαμόρφωση φουρκέτας. Για να διαπιστωθεί αυτό, τοποθετήθηκε στο πλασμίδιο αλληλουχία AT(12) επαναλήψεων που είναι το ίδιο ευάλωτο να σχηματίσει δομή φουρκέτας αλλά όχι Z-DNA. Βρέθηκε ότι έχει πολύ μικρότερη συχνότητα μεταλλάξεων σε σύγκριση με αλληλουχία CG(12) επαναλήψεων. [110] Στα κύτταρα των θηλαστικών οι CG (14) επαναλήψεις προκαλούν σπάσιμο στην έλικα του DNA γύρω από την Z-διαμόρφωση και αυτό οδηγεί σε διαγραφή ολόκληρων των επαναλήψεων. [111]

Επίσης, εναλλαγές πουρίνης και πυριμιδίνης ικανές για να σχηματίσουν Z-διαμορφώσεις είναι στόχοι καθαρισμού για το ένζυμο τοποϊσομεράση II στα σπονδυλωτά *in vivo*. Η καθαρή τοποϊσομεράση II από όρνιθα καθαρίζει εναλλαγές πουρίνης-πυριμιδίνης (GT, AT, GC) *in vitro*. [112]



## H – DNA

H – DNA (ή ενδομοριακή τρίκλωνη έλικα DNA). Μία από τις αλυσίδες της πολυπυριμινικής – πολυπυριμιδινικής περιοχής που περιέχει συμμετρία καθρέφτη αναδιπλώνεται και σχηματίζει δεσμούς υδρογόνου Hoogsten με τα κατάλοιπα πουρίνης της διπλής έλικας του DNA. Η δομή που προκύπτει είναι στη μισή περιοχή τριπλή έλικα του DNA και στην άλλη μισή μονή έλικα. [99]



**Εικόνα 8.** Σχηματική απεικόνιση της H – DNA διαμόρφωσης.

Αλληλουχίες που μπορούν να πάρουν H-DNA διαμόρφωση συναντώνται φυσιολογικά στο γονιδίωμα του ανθρώπου μία φορά κάθε 50.000 βάσεις.  
[113]

Τα σπασίματα στο DNA μπορεί να οφείλονται στην καθυστέρηση της DNA πολυμεράσης κατά τον διπλασιασμό του DNA σ' αυτή την ασυνήθιστη διαμόρφωση. GA και GAA επαναλήψεις μπορούν να προκαλέσουν προσωρινή διακοπή ή και μόνιμη διακοπή της αντιγραφής του DNA στα θηλαστικά. [114] [115] Πάντως πρέπει να ληφθεί υπόψη ότι η ίδια δομή του DNA έχει διαφορετική τύχη σε διαφορετικά είδη κυττάρων. [99]

### Σιωπηρές γενετικές μεταλλάξεις του *p53*

Είναι σκόπιμο να διερευνήσουμε τις σιωπηρές μεταλλάξεις του γονιδίου *p53*, διότι δεν δίνουν πλεονέκτημα ή μειονέκτημα στα κύτταρα. Η παραγομένη πρωτεΐνη είναι η ίδια. Επομένως όλα τα κύτταρα που έχουν σιωπηρή μετάλλαξη θα τη διατηρήσουν δίχως επιλογή από το περιβάλλον. Δεν θα πολλαπλασιαστούν ευνοϊκότερα από τα άλλα κύτταρα αλλά ούτε έχουν μειονέκτημα. Αντίθετα κύτταρα με αλλοιώσεις της πρωτεΐνης TP53 επηρεάζονται από την πίεση του περιβάλλοντος. Έτσι ορισμένες μεταλλάξεις γονιδίων που προσδίδουν μειονέκτημα στα κύτταρα θα εξαφανιστούν μαζί με τους φορείς τους. Όταν θα μελετηθούν τα κύτταρα στο εργαστήριο θα έχουν απομείνει τα υπόλοιπα κύτταρα και δε θα γνωρίζουμε τελικά

τη συχνότητα των μειονεκτικών μεταλλάξεων. Αντίθετα, τα κύτταρα που προκαλούν καρκινικό φαινότυπο θα μελετηθούν εκτενέστερα στο εργαστήριο καθώς οι φορείς τους θα νοσήσουν και θα αναζητήσουν ιατρική φροντίδα. Έτσι θα βρεθούν με μεγαλύτερη συχνότητα σε σχέση με τον γενικό πληθυσμό. Επίσης, για να βρεθεί η συχνότητα αλλά και η ποσότητα των σιωπηρών μεταλλάξεων θα πρέπει να γίνει έλεγχος του γενικού πληθυσμού δίχως κανένα σύμπτωμα ούτε και σε κυτταρικό επίπεδο, πράγμα το οποίο είναι πολύ δύσκολο.

Από την ηλεκτρονική διεύθυνση του Διεθνούς Οργανισμού για την Έρευνα του Καρκίνου (International Agency for Research on Cancer) [www-p53.iarc.fr](http://www-p53.iarc.fr) και τη βάση δεδομένων [TP53GermlineRefR10](#) βρέθηκαν τέσσερις βιβλιογραφικές αναφορές σε σιωπηρές p53 γενετικές μεταλλάξεις. Δημοσιευμένες μέσω internet είναι τρεις από αυτές. Προσεκτική τους μελέτη δείχνει ότι και οι τρεις, παρόλο που ερεύνησαν διαφορετικές οικογένειες, περιγράφουν μετάλλαξη στο κωδικίο 125 με το ίδιο αποτέλεσμα.

Στην πρώτη περίπτωση περιγράφεται σιωπηρή μετάλλαξη του κωδικίου 125 ACG>ACC που έχει ως αποτέλεσμα τη διατήρηση 109 βάσεων από το ιντρόνιο 4. Παρόλο δηλαδή που το αμινοξύ που κωδικοποιεί το κωδικίο 125 δεν αλλάζει, εντούτοις αλλάζει σημειακά η αλληλουχία της περιοχής και οι μηχανισμοί που αποκόβουν τα ενδόνια δεν μπόρεσαν να λειτουργήσουν. Έτσι το μεταγράφημα είναι κατά 109 βάσεις μακρύτερο. [116]

Στη δεύτερη περίπτωση αναφέρεται ξανά η μετάλλαξη της γενετικής σειράς στη θέση ματίσματος του ιντρονίου 4 . Στο κωδικίο 125 μετατρέπεται ACG>ACA, με αποτέλεσμα να διατηρείται μέρος του ιντρονίου 4 στο τελικό μεταγράφημα.

*Table 2* 5' splice donor sequences

Sequence <sup>a</sup>	Exon/Intron -/+ 321/123456 <sup>b</sup>	Bases matching consensus sequence
Consensus sequence	5' $\begin{matrix} \text{C} \\ \text{A} \end{matrix} \text{AG/GT} \begin{matrix} \text{A} \\ \text{G} \end{matrix} \text{AGT} \text{ 3'}$	
WT p53 intron 4 sequence	ACG/GTCAGT	7/9
Mut p53 intron 4 sequence	ACA/GTCAGT <sup>c</sup>	6/9

<sup>a</sup> WT, wild-type; Mut, mutant.

<sup>b</sup> The position numbers of the bases in the consensus sequence are given.

<sup>c</sup> The guanine to adenine transition in the splice sequence is in bold-face type.



**Εικόνα 9.** Σύγκριση των αλληλουχιών consensus, άγριου τύπου, μεταλλαγμένου τύπου του γονιδίου *p53* στην περιοχή ματίσματος του εξωνίου 4.

Η consensus αλληλουχία είναι 9 βάσεις AAG/GTAAGT και η wild type αλληλουχία του *p53* στη θέση ματίσματος είναι ACG/GTCAGT. Δηλαδή είναι όμοιες μόνο 7 από τις 9 βάσεις και αυτές είναι αρκετές για να γίνει σωστά το splicing. Όμως όταν συμβαίνει και ακόμη μία μετάλλαξη G>A τότε η ομοιότητα μεταξύ consensus αλληλουχίας και αλληλουχίας του *p53* στη θέση ματίσματος διαφοροποιείται κατά 6 βάσεις στο σύνολο των 9 βάσεων. Αυτό είναι αρκετό για να εμποδίσει το μάτισμα σ' αυτή τη θέση. [117]

Στην τρίτη βιβλιογραφική αναφορά περιγράφεται μετάλλαξη στο κωδικό 125 του εξωνίου 4 ACG/GT>ACA/GT που επηρεάζει το μάτισμα και οδηγεί στην παραγωγή μεγαλύτερου *p53* μεταγραφήματος από το φυσιολογικό. [118]

Επομένως οι περιγραφόμενες μεταλλάξεις μπορεί να είναι σιωπηρές ως προς την τριπλέτα του γενετικού κώδικα, προκαλούν όμως αλλαγές στην παραγόμενη πρωτεΐνη. Έτσι χάνουν την ιδιαιτερότητά τους και η συχνότητά τους επηρεάζεται όχι μόνο από τη μεταλλαξιμότητα των προτύπων – νουκλεοτιδίων, αλλά και από τη φυσική επιλογή, που επιλέγει στα καρκινικά κύτταρα μεταλλάξεις που έχουν αρνητικό βιολογικό αποτέλεσμα στη λειτουργία του *p53*.

### **Γενετικά σύνδρομα που οφείλονται σε μεταλλάξεις του *p53***

Οι γενετικές μεταλλάξεις του *p53* έχουν ταυτοποιηθεί κυρίως σε οικογένειες με κυρίαρχο και κληρονομικό Li-Fraumeni σύνδρομο. Αυτά τα άτομα είναι ευάλωτα σε διάφορα είδη νεοπλασμάτων από πολύ μικρή ηλικία. Τα κριτήρια του συνδρόμου είναι η ύπαρξη προγόνου με σάρκωμα που εμφανίστηκε σε ηλικία κάτω των 45 ετών, ο οποίος είχε άλλους δύο συγγενείς που ανέπτυξαν καρκίνο σε ηλικία κάτω των 45 ετών. Ο ένας συγγενής του προγόνου είναι πρώτου βαθμού και ο άλλος πρώτου ή δεύτερου βαθμού. Ασθενείς με LFS είναι ιδιαίτερα ευάλωτοι σε νεοπλασμάτα του μαστού, των επινεφριδίων, σαρκώματα μαλακών μορίων και οστών, οξείες λευχαιμίες και όγκους του εγκεφάλου. Η καλύτερη κατανόηση του συνδρόμου βοηθά στη γενετική συμβουλή, πρωιμότερη διάγνωση και αποτελεσματικότερη θεραπεία. [119] [120] [121] [122] [123] [124] [125] [126]

Υπάρχουν περιπτώσεις που δεν εκπληρώνουν τα κριτήρια του συνδρόμου Li – Fraumeni, διότι ο πρόγονος τους είχε δύο συγγενείς που ανέπτυξαν καρκίνο σε ηλικία μεγαλύτερη των 45 ετών. Τότε ονομάζουμε το σύνδρομο αυτό Li – Fraumeni Like (LFL). Σε πολλές

περιπτώσεις κληρονομικού καρκίνου δεν εκπληρώνονται τα παραπάνω κριτήρια και τις χαρακτηρίζουμε ως Family History (Οικογενειακό ιστορικό). Αν δεν υπάρχει οικογενειακό ιστορικό καρκινογένεσης τότε τα άτομα ταξινομούνται ως No Family History (Δεν υπάρχει οικογενειακό ιστορικό).

## II. ΥΛΙΚΟ – ΜΕΘΟΔΟΙ

Το υλικό της εργασίας συλλέχθηκε από την ιστοσελίδα του Διεθνούς Οργανισμού Έρευνας του Καρκίνου (International Agency for Research on Cancer) που βρίσκεται στο διαδίκτυο στη σελίδα ([www.iarc.fr](http://www.iarc.fr)). Μελετήσαμε αποκλειστικά το γονίδιο *p53*. Οι βάσεις δεδομένων που χρησιμοποιήθηκαν βρίσκονται στην ηλεκτρονική διεύθυνση <http://www-p53.iarc.fr> και είναι δύο: οι γενετικές μεταλλάξεις του γονιδίου *p53* και οι σωματικές. Τα δεδομένα αυτών των βάσεων περιλαμβάνουν 21.587 καταγραφές για τα σωματικά κύτταρα και 1.424 εγγραφές για τα γεννητικά. Η τελευταία ενημέρωση των βάσεων έγινε τον Ιούνιο του 2005.

### Προγράμματα ανάλυσης των βάσεων δεδομένων

Για την επεξεργασία και ανάλυση των βάσεων δεδομένων και τη δημιουργία των γραφημάτων χρησιμοποιήθηκε πρόγραμμα διαχείρισης spreadsheet. Επίσης επιπλέον ανάλυση του φάσματος των μεταλλάξεων στα γενετικά και στα σωματικά κύτταρα έγινε με τη βοήθεια του iMARS, που είναι εξειδικευμένο πρόγραμμα στην μαθηματική επεξεργασία των μεταλλάξεων. Το iMARS αναπτύχθηκε ως πρόγραμμα οπτικοποίησης και επεξεργασίας των δεδομένων μεταλλάξεων σε κάθε αλληλουχία DNA. Μπορεί να λειτουργήσει σε λειτουργικό Microsoft Windows και είναι διαθέσιμο στη διεύθυνση <http://www.clinical-bioinformatics.net/imars/imars.zip>. Έπειτα από αίτηση στους δημιουργούς του προγράμματος, επιτράπηκε η χρήση του για τη συγκεκριμένη εργασία. Λόγω του τρόπου εισαγωγής των δεδομένων δεν είναι δυνατόν να εισαχθούν δεδομένα στην τριπλέτα 0, δηλαδή σε ιντρόνιο. Επομένως αυτές οι μεταλλάξεις δεν εισήχθησαν. Επίσης δεν κατέστη δυνατό να εισαχθούν οι κατηγορίες μεταλλάξεων CC tandem, complex, NA και tandem επειδή το πρόγραμμα δεν έχει προβλέψει τέτοιου είδους δεδομένα. [127]

Η εισαγωγή των γενετικών μεταλλάξεων έγινε εισάγοντας μία μετάλλαξη ανά οικογένεια, για να μην επηρεαστεί η συχνότητα των μεταλλάξεων από τον αριθμό των απογόνων. Εισήχθησαν 258 εγγραφές και δεν εισήχθησαν 24 εγγραφές ιντρονίων με τον αριθμό 0. Όλες οι εγγραφές των γενετικών κυττάρων ανά οικογένεια είναι 282 ( $258 + 24 = 282$ ), επομένως εισήχθησαν όλες οι μεταλλάξεις που μπορούν να επηρεάσουν τα συμπεράσματά μας στη μελέτη των εξωνίων. Εισήχθησαν 20.706 εγγραφές στα σωματικά κύτταρα από το σύνολο των 21.587. Οι μεταλλάξεις στην τριπλέτα 0 είναι 620 και αυτές δεν επηρεάζουν την εξαγωγή των συμπερασμάτων για τα εξώνια. Επομένως δεν εισήχθησαν  $21.587 - (20.706 + 620) = 261$  μεταλλάξεις που ανήκουν στην κατηγορία CC tandem, complex, NA και tandem και επηρεάζουν την ανάλυση των εξωνίων.

### Στατιστική επεξεργασία

Για τη στατιστική επεξεργασία χρησιμοποιήθηκε η δοκιμασία  $\chi^2$  και ο συντελεστής συσχέτισης του Pearson. Επίσης χρησιμοποιήθηκαν και οι στατιστικές αναλύσεις του προγράμματος iMARS. [127]

Για τον υπολογισμό των διαφορών μεταξύ των παρατηρούμενων και των αναμενόμενων αντικαταστάσεων νουκλεοτιδίων στα ζεύγη G:C και A:T και της στατιστικής τους σημαντικότητας χρησιμοποιήθηκε ο τύπος:

$\chi^2 = \sum (\text{παρατηρούμενων} - \text{αναμενόμενων})^2 / \text{αναμενόμενων}$ , με έναν βαθμό ελευθερίας. Και στη συνέχεια έγινε η δοκιμασία  $\chi^2$ .

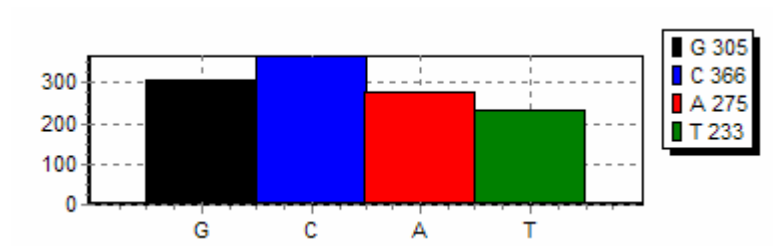
Για τον υπολογισμό της σχετικής μεταλλαξιμότητας των δινουκλεοτιδίων χρησιμοποιήθηκε ο τύπος των Cooper and Krawczak. Ο τύπος υπολογισμού για κάθε δινουκλεοτίδιο χωριστά είναι  $R_{dm} = (O/E) * M$ , όπου  $R_{dm}$  είναι η σχετική μεταλλαξιμότητα,  $O$  είναι ο αριθμός των παρατηρούμενων μεταλλάξεων του δινουκλεοτιδίου,  $E$  είναι οι αναμενόμενες μεταλλάξεις, και  $M$  είναι το μικρότερο κλάσμα  $O/E$  που καταγράφηκε από τους υπολογισμούς για όλα τα δινουκλεοτίδια. [128]

Τα hotspots των μεταλλάξεων υπολογίστηκαν από το τεστ του Tarone στο πρόγραμμα iMARS. [129]

Η σύγκριση δύο φασμάτων μεταλλάξεων έγινε με τη βοήθεια του υπεργεωμετρικού τεστ, στο οποίο επισέρχεται και η διόρθωση του Bonferroni. [130] [131]

### III. ΑΠΟΤΕΛΕΣΜΑΤΑ

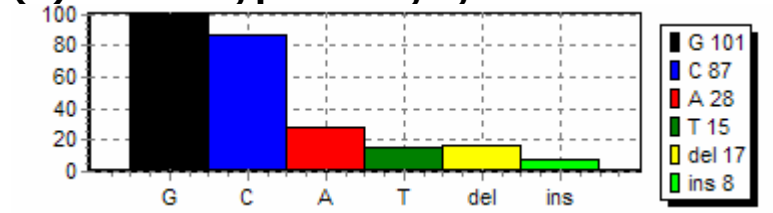
**1. Από την ανάλυση της αλληλουχίας του γονιδίου p53 και του τύπου των μεταλλάξεων που έχουν καταγραφεί, με τη βοήθεια του προγράμματος iMARS, προκύπτουν τα εξής αποτελέσματα:**



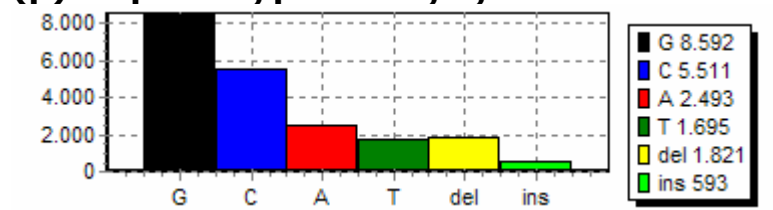
**Σχήμα 1. Ποσοστά των τεσσάρων νουκλεοτιδίων στο γονίδιο του p53**

Το μήκος της αλληλουχίας είναι 1179 νουκλεοτίδια και το GC περιεχόμενο 56,91%. Το γονίδιο περιέχει περισσότερα νουκλεοτίδια κυτοσίνης.

#### (α) Γενετικές μεταλλάξεις



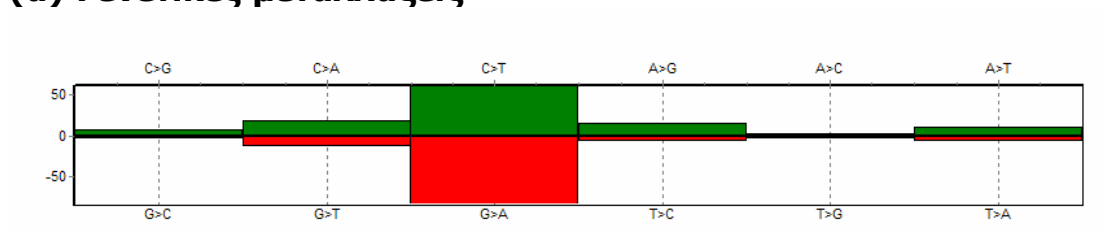
#### (β) Σωματικές μεταλλάξεις



**Σχήμα 2. Μεταλλάξεις των τεσσάρων νουκλεοτιδίων.**

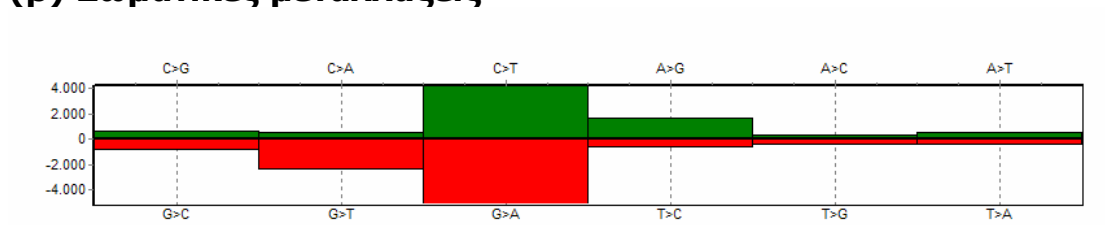
(α) Μεταλλάξεις κάθε νουκλεοτιδίου συνολικά για τα γενετικά κύτταρα. (β) Μεταλλάξεις κάθε νουκλεοτιδίου συνολικά για τα σωματικά κύτταρα.

Παρατηρούμε ότι οι μεταλλάξεις κυτοσίνης είναι περισσότερες στις γενετικές μεταλλάξεις έναντι των σωματικών στατιστικώς πολύ σημαντικά με  $p=0,0000$ .

**(α) Γενετικές μεταλλάξεις**

Base	OBS	EXP	CHI 2	P
C	87	102,55	5,18457	0,02
G	101	85,45		

Base	OBS	EXP	CHI 2	P
A	28	23,28	2,92318	0,09
T	15	21,5		

**(β) Σωματικές μεταλλάξεις**

Base	OBS	EXP	CHI 2	P
C	5511	7671,53	1334,255	0,00
G	8592	6431,47		

Base	OBS	EXP	CHI 2	P
A	2493	2267,13	98,53107	0,00
T	1695	2094		

**Σχήμα 3. Από τα ζεύγη νουκλεοτιδίων G:C και A:T διερευνάται ποιο νουκλεοτίδιο μεταλλάσσεται περισσότερο σε κάθε τύπο μετάλλαξης.**

Επομένως στο γονίδιο *p53* και στις γενετικές και στις σωματικές μεταλλάξεις παρατηρούνται ξεκάθαρα περισσότερες μεταλλάξεις γουανίνης και αδενίνης από τις αναμενόμενες και λιγότερες μεταλλάξεις κυτοσίνης και θυμίνης από τις αναμενόμενες.

Παρατηρούμε ότι υπερέχουν οι μεταλλάξεις της γουανίνης έναντι της κυτοσίνης στην κωδική αλυσίδα. Αυτό σημαίνει ότι μεταλλάσσεται περισσότερο απ' όλα η γουανίνη της κωδικής αλυσίδας ή πιθανόν και η κυτοσίνη της συμπληρωματικής αλυσίδας.

Άρα στα γενετικά κύτταρα η γουανίνη μεταλλάσσεται στατιστικώς πολύ σημαντικά περισσότερο από τη κυτοσίνη με πιθανότητα  $p=0,02$ . Μπορεί όμως θεωρητικά να συμβαίνουν μεταλλάξεις μεταξύ γουανίνης και κυτοσίνης με την ίδια συχνότητα αλλά οι μεταλλάξεις της γουανίνης να οδηγούν συχνότερα σε καρκινική εξαλλαγή και έτσι να έχουν περισσότερες εγγραφές στις βάσεις δεδομένων.

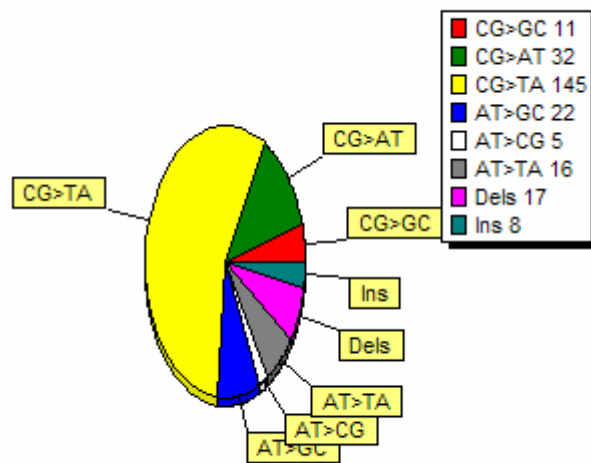
Γι αυτό θεωρούμε πιθανούς παράγοντες του φαινομένου τα εξής:

- ✓ Η γουανίνη ως βάση είναι πιο επηρεπής σε μεταλλάξεις από την κυτοσίνη
- ✓ Το κύτταρο έχει πιο αποτελεσματικούς μηχανισμούς επιδιόρθωσης των μεταλλάξεων της κυτοσίνης
- ✓ Επιλέγονται οι μεταλλάξεις της γουανίνης

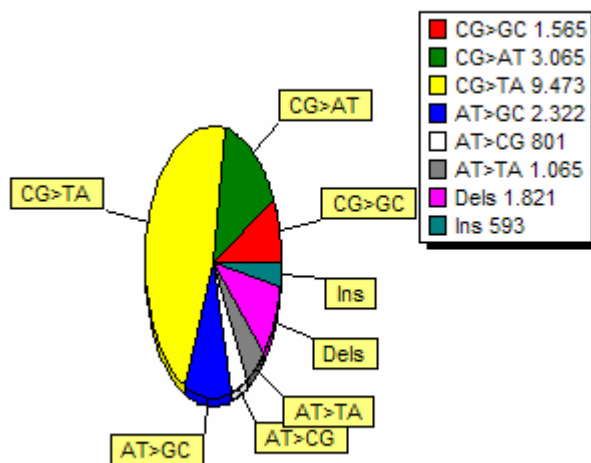
Τα ίδια συμπεράσματα ισχύουν και για το ζεύγος αδενίνης – θυμίνης μόνο που τα αποτελέσματα είναι λιγότερο στατιστικώς σημαντικά.

Στα σωματικά κύτταρα ισχύουν τα συμπεράσματα που διατυπώθηκαν για τα γενετικά και όλα τα αποτελέσματα είναι στατιστικώς πολύ σημαντικά με  $p=0,00$ .

Ακόμη, στα γενετικά κύτταρα υπερέχουν οι μεταλλάξεις C>A έναντι των σωματικών. Αντίθετα στα σωματικά κύτταρα υπερέχουν οι μεταλλάξεις G>C, G>T και G>A έναντι των γενετικών.

**(α) Γενετικές μεταλλάξεις**

Τύπος μετάλλαξης	%	Απόλυτη τιμή
CG>GC	11	4,3
CG>AT	32	12,5
CG>TA	145	56,64
AT>GC	22	8,59
AT>CG	5	1,95
AT>TA	16	6,25
Dels	17	6,64
Ins	8	3,13

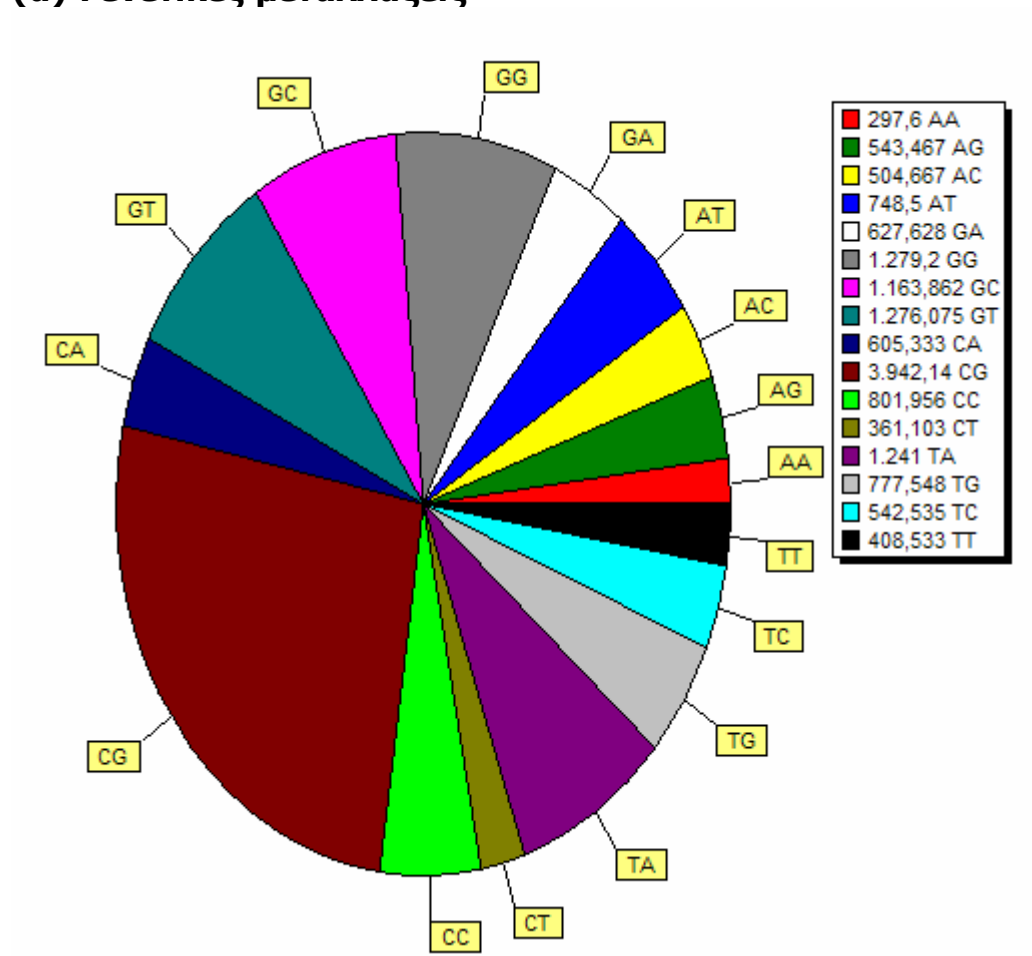
**(β) Σωματικές μεταλλάξεις**

Τύπος μετάλλαξης	%	Απόλυτη τιμή
CG>GC	1565	7,56
CG>AT	3065	14,8
CG>TA	9473	45,75
AT>GC	2322	11,21
AT>CG	801	3,87
AT>TA	1065	5,14
Dels	1821	8,79
Ins	593	2,86

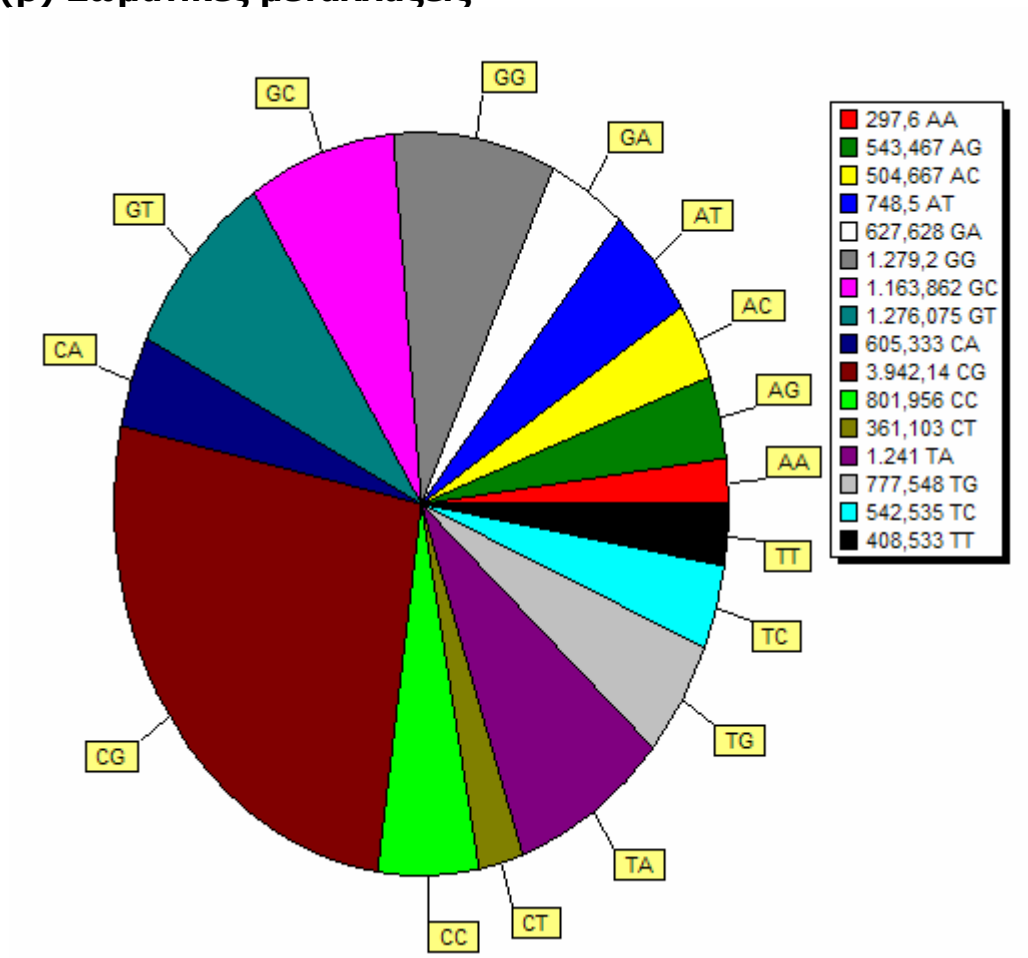
**Σχήμα 4. Συχνότητα τύπου μετάλλαξης κάθε δινουκλεοτιδίου**



## (α) Γενετικές μεταλλάξεις



Dinuc	Number	Frequency	Mutations	Expected	RDM	LH	%
AA	65	0,055	14	28,472	5,169	0,026	2,6
AG	90	0,076	16	39,423	4,267	0,021	2,1
AC	72	0,061	12	31,538	4	0,02	2
AT	48	0,041	12	21,025	6	0,03	3
GA	86	0,073	15	37,671	4,186	0,021	2,1
GG	80	0,068	40	35,042	12	0,06	6
GC	86	0,073	77	37,671	21,488	0,107	10,7
GT	53	0,045	24	23,216	10,868	0,054	5,4
CA	99	0,084	23	43,365	5,576	0,028	2,8
CG	42	0,036	146	18,397	83,429	0,415	41,5
CC	137	0,116	57	60,01	9,985	0,05	5
CT	87	0,074	12	38,109	3,31	0,016	1,6
TA	24	0,02	13	10,513	13	0,065	6,5
TG	93	0,079	28	40,737	7,226	0,036	3,6
TC	71	0,06	21	31,1	7,099	0,035	3,5
TT	45	0,038	6	19,711	3,2	0,016	1,6
	1178	1	516	516	200,8027	1	

**(β) Σωματικές μεταλλάξεις**

Dinuc	Number	Frequency	Mutations	Expected	RDM	LH	%
AA	65	0,055	806	2284,932	297,6	0,02	2
AG	90	0,076	2038	3163,752	543,467	0,036	3,6
AC	72	0,061	1514	2531,002	504,667	0,033	3,3
AT	48	0,041	1497	1687,334	748,5	0,05	5
GA	86	0,073	2249	3023,141	627,628	0,042	4,2
GG	80	0,068	4264	2812,224	1279,2	0,085	8,5
GC	87	0,074	4219	3058,294	1163,862	0,077	7,7
GT	53	0,045	2818	1863,098	1276,075	0,084	8,4
CA	99	0,084	2497	3480,127	605,333	0,04	4
CG	43	0,037	7063	1511,57	3942,14	0,261	26,1
CC	135	0,115	4511	4745,628	801,956	0,053	5,3
CT	87	0,074	1309	3058,294	361,103	0,024	2,4
TA	24	0,02	1241	843,667	1241	0,082	8,2
TG	93	0,079	3013	3269,211	777,548	0,051	5,1
TC	71	0,06	1605	2495,849	542,535	0,036	3,6
TT	45	0,038	766	1581,876	408,533	0,027	2,7
	1178	1	41410	41410	15121,15	1	

**Σχήμα 5. Κατανομή των γενετικών και σωματικών μεταλλάξεων ανά δινουκλεοτίδιο**

Σύμφωνα με τη σχετική μεταλλαξιμότητα που υπολογίσαμε στο τελευταίο κελί ως ποσοστό σε σχέση με τη μικρότερη μεταλλαξιμότητα, μπορούμε να βάλουμε σε σειρά τα δινουκλεοτίδια αρχίζοντας με αυτό που έχει τη μεγαλύτερη μεταλλαξιμότητα.

(α) Γενετικές μεταλλάξεις

**CG > GC > TA > GG > GT > CC > TG > TC > AT > CA > AA > GA = AG > AC > TT = CT**

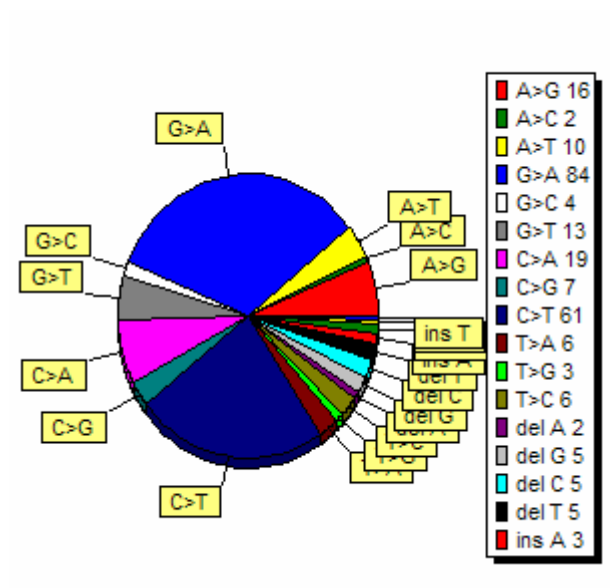
(β) Σωματικές μεταλλάξεις

**CG > GG > GT > TA > GC > CC > TG > AT > GA > CA > TC = AG > AC > TT > CT > AA**

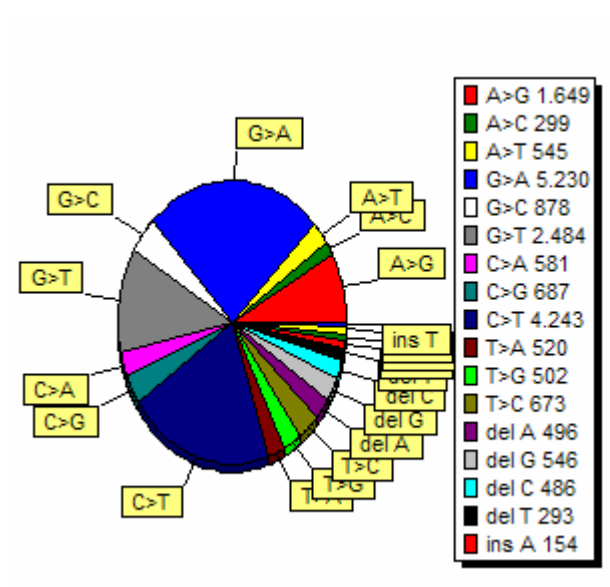
Πιο μεταλλασσόμενο νουκλεοτίδιο είναι το CG. Λιγότερο μεταλλασσόμενο το CT. Είναι αξιοσημείωτο ότι στις πρώτες θέσεις είναι το δινουκλεοτίδιο TA. Αυτό σημαίνει ότι δεν είναι απαραίτητη προϋπόθεση για μετάλλαξη η περιεκτικότητα σε γουανίνη και κυτοσίνη του δινουκλεοτιδίου. Ο αριθμός των παρατηρούμενων μεταλλάξεων των δινουκλεοτιδίων είναι πιθανότατα συνάρτηση της μεταλλαξιμότητάς τους και της τοξικότητάς τους στη λειτουργία της πρωτεΐνης.

Δινουκλεοτίδια	Συχνότητα % γενετικών μεταλλάξεων	Συχνότητα % σωματικών μεταλλάξεων	Δοκιμασία $\chi^2$
AA	2,6	2	0,7821
AG	2,1	3,6	0,5356
AC	2	3,3	0,5773
AT	3	5	0,4881
GA	2,1	4,2	0,4100
GG	6	8,5	0,5261
GC	10,7	7,7	0,5033
GT	5,4	8,4	0,4347
CA	2,8	4	0,6509
CG	41,5	26,1	0,1048
CC	5	5,3	0,9274
CT	1,6	2,4	0,6921
TA	6,5	8,2	0,6687
TG	3,6	5,1	0,6186
TC	3,5	3,6	0,9706
TT	1,6	2,7	0,5997

Παρατηρούμε ότι υπάρχουν διαφορές μεταξύ γενετικών και σωματικών μεταλλάξεων. Αυτές οι διαφορές όμως δεν είναι στατιστικώς σημαντικές, όπως προκύπτει από τη δοκιμασία  $\chi^2$ . Είναι όμως αξιοσημείωτη η τιμή  $p=0,1048$  στη σύγκριση του δινουκλεοτιδίου CG.

**(α) Γενετικές μεταλλάξεις**

	N	%
A>G	16	6,25
A>C	2	0,78
A>T	10	3,91
G>A	84	32,81
G>C	4	1,56
G>T	13	5,08
C>A	19	7,42
C>G	7	2,73
C>T	61	23,83
T>A	6	2,34
T>G	3	1,17
T>C	6	2,34
del A	2	0,78
del G	5	1,95
del C	5	1,95
del T	5	1,95
ins A	3	1,17
ins G	3	1,17
ins C	1	0,39
ins T	1	0,39

**(β) Σωματικές μεταλλάξεις**

	N	%
A>G	1649	7,96
A>C	299	1,44
A>T	545	2,63
G>A	5230	25,26
G>C	878	4,24
G>T	2484	12
C>A	581	2,81
C>G	687	3,32
C>T	4243	20,49
T>A	520	2,51
T>G	502	2,42
T>C	673	3,25
del A	496	2,4
del G	546	2,64
del C	486	2,35
del T	293	1,42
ins A	154	0,74
ins G	181	0,87
ins C	154	0,74
ins T	104	0,5

**Σχήμα 6. Συχνότητες μεταλλάξεων νουκλεοτιδίων.**

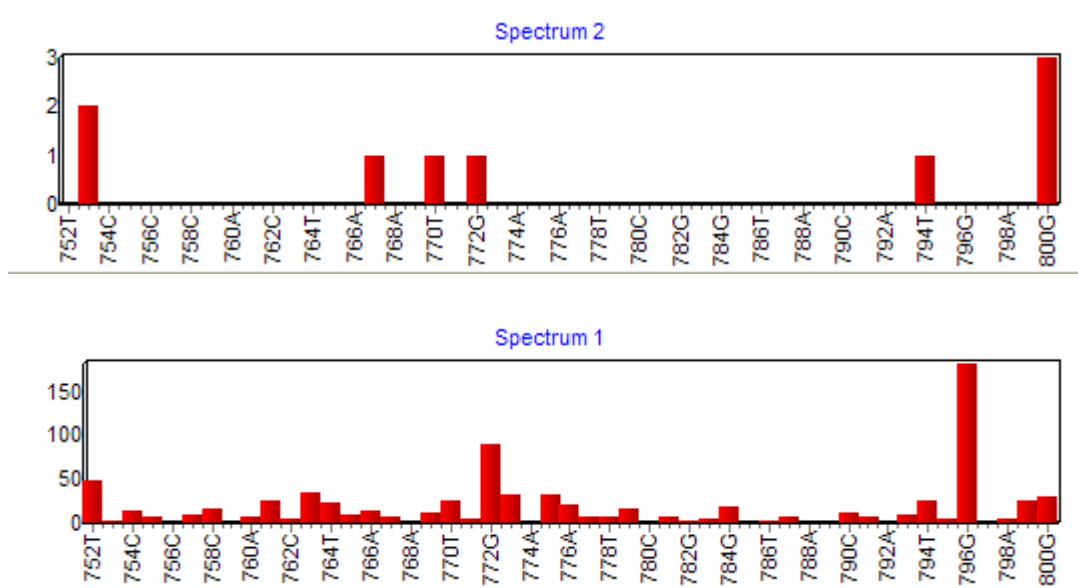
Όπως φαίνεται, οι συχνότητες μεταξύ γενετικών και σωματικών μεταλλάξεων διαφέρουν. Συγκριτικά μεγαλύτερες διαφορές παρατηρούνται στις μεταλλάξεις G>A, G>T, G>C και C>A.

	(α) Γενετικές μεταλλάξεις	(β) Σωματικές μεταλλάξεις	Δοκιμασία $\chi^2$
<b>G&gt;A</b>	32,81	25,26	0,382894
<b>G&gt;C</b>	1,56	4,24	0,272594
<b>G&gt;T</b>	5,08	12	0,107834
<b>C&gt;A</b>	7,42	2,81	0,159676

Με τη δοκιμασία  $\chi^2$  δεν προκύπτουν στατιστικώς σημαντικά συμπεράσματα από τη σύγκριση γενετικών και σωματικών μεταλλάξεων. Φτάνουν όμως οι μεταλλάξεις G>T το επίπεδο σημαντικότητας του 10%. Οι μεταλλάξεις αυτές θεωρείται ότι προκαλούνται από μεθυλίωση της γουανίνης στα σωματικά κύτταρα. [62]

### **Συσχέτιση της μεταλλαξιμότητας ορισμένης αλληλουχίας και της περιεκτικότητας σε G:C με τη βοήθεια του υπεργεωμετρικού τεστ του Monte Carlo.**

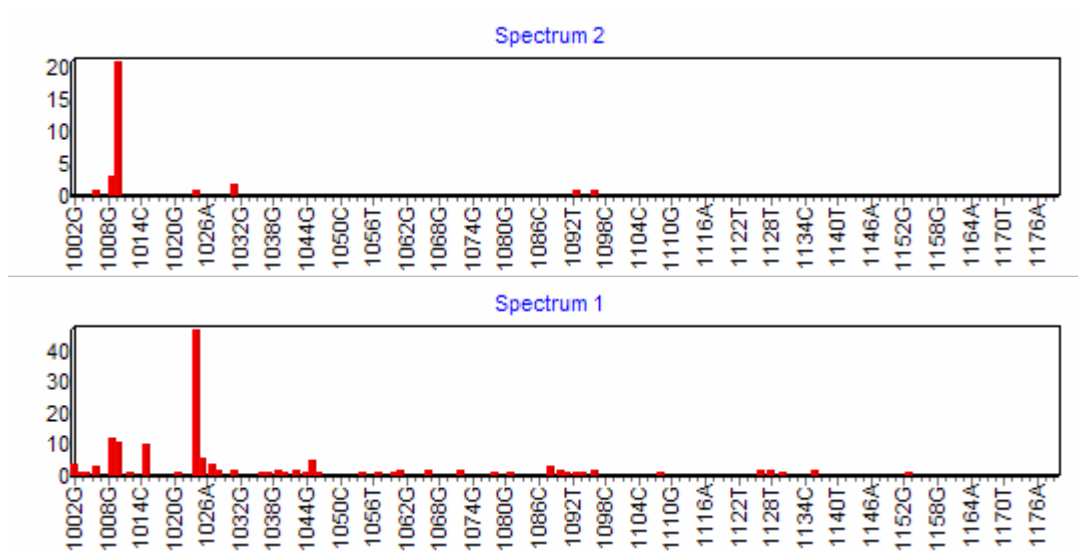
Το φάσμα 2 ανήκει στις γενετικές μεταλλάξεις και το φάσμα 1 στις σωματικές. Η ανάλυση των φασμάτων έγινε ανά 50 βάσεις λόγω περιορισμών του προγράμματος, εκτός από τις βάσεις 1000-1179. Επίσης λόγω περιορισμών του προγράμματος δεν είναι εφικτή η σύγκριση του φάσματος μεταλλάξεων μεταξύ 450-550 και 800-900. Από τα αποτελέσματα που ήταν εφικτό να εξαχθούν παρατηρούμε ότι τα δύο φάσματα μεταλλάξεων έχουν στατιστικώς σημαντικές διαφορές στις περιοχές 750-800 και 1000-1179. Οι παραπάνω αριθμοί αντιστοιχούν στους αριθμούς που έχουν τα νουκλεοτίδια στην αλληλουχία του p53.



**Σχήμα 7 (α). Monte Carlo hypergeometric test για τη σύγκριση των φασμάτων μεταλλάξεων των νουκλεοτιδίων 752-800.**

Η πιθανότητα ότι τα δύο φάσματα μεταλλάξεων είναι ίδια είναι: 0,004118

p value 95%, όρια εμπιστοσύνης: 0,001074, 0,007



**Σχήμα 7 (β). Monte Carlo hypergeometric test για τη σύγκριση των φασμάτων μεταλλάξεων των νουκλεοτιδίων 1002-1176.**

Η πιθανότητα ότι τα δύο φάσματα μεταλλάξεων είναι ίδια είναι: 0,000000

Οι στατιστικώς σημαντικές διαφορές βρίσκονται στα κωδικία 251-267 και 334-393.

Παρατηρούμε ότι στην πρώτη περίπτωση κάνουν τη διαφορά η μετάλλαξη του νουκλεοτιδίου 753C (κωδικό 251) και 800G (κωδικό 267) στα γενετικά, η 772G (κωδικό 258) και 796G (κωδικό 266) στα σωματικά.

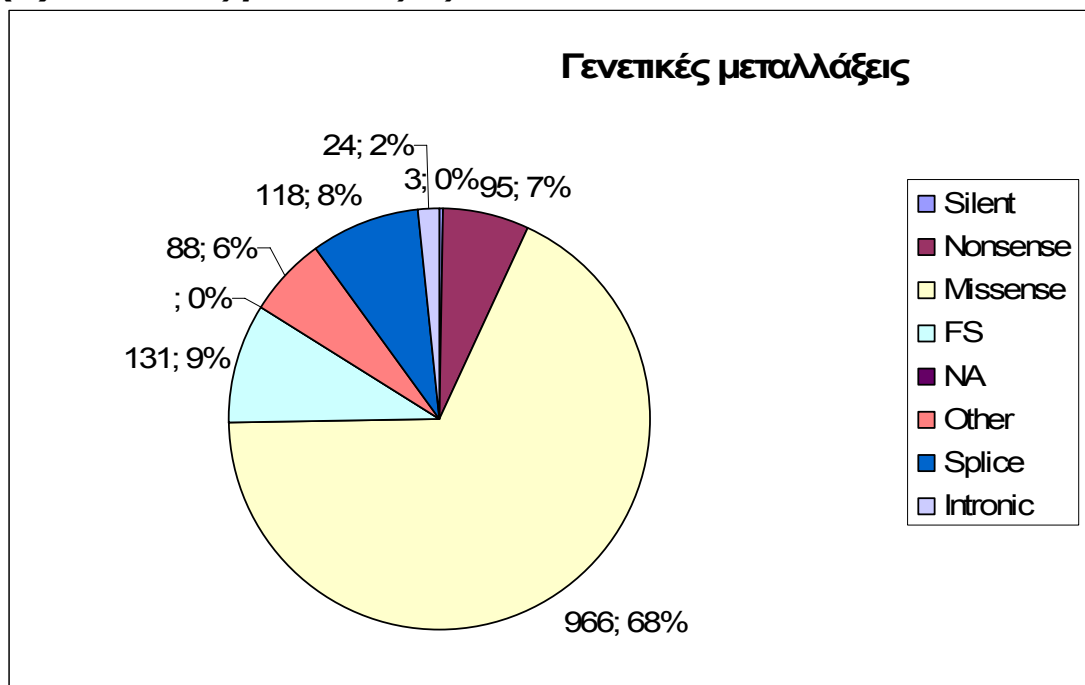
Είναι αξιοσημείωτο ότι τα κωδικία που μεταβάλλονται στα γενετικά (251 και 267) βρίσκονται σε πολύ πλούσια G C αλυσίδα. Συνολικά 10 από τα 15 νουκλεοτίδια που συνιστούν τα κωδικία 248-253 και 265-269 είναι GC. Στα σωματικά κύτταρα αυτό το ποσό πέφτει στα 8 από τα 15 νουκλεοτίδια στα κωδικία 256-260 και στα 9 νουκλεοτίδια στα κωδικία 264-268.

Άρα, πρώτες γίνονται οι μεταλλάξεις σε περιοχές με πολύ υψηλό ποσό G:C (γενετικές) και στη συνέχεια γίνονται μεταλλάξεις λόγω επιγενετικών τροποποιήσεων (σωματικές) σε περιοχές με χαμηλότερο ποσό G:C. Επίσης η τριπλέτα 267 περιέχει θέση CpG η οποία παρουσιάζει αυξημένη μεταλλαξιμότητα στα γενετικά κύτταρα, όχι όμως και στα σωματικά.

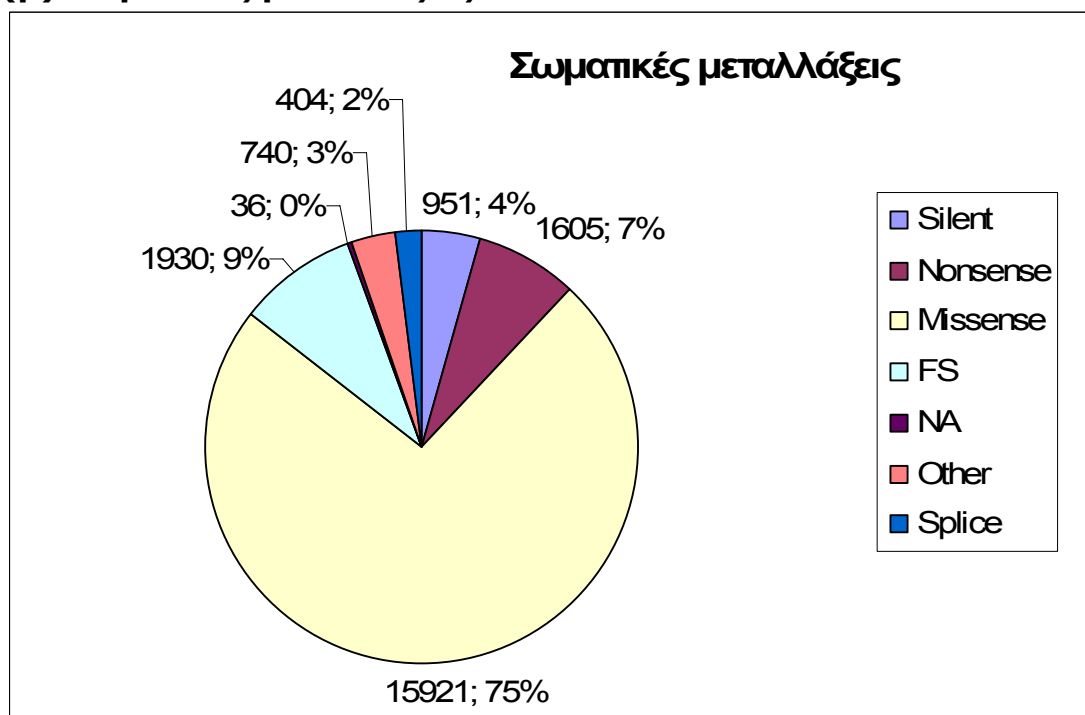
Στη δεύτερη σύγκριση των φασμάτων παρατηρούμε ότι οι διαφορές βρίσκονται στο νουκλεοτίδιο 1010G στα γενετικά και στο 1024C στα σωματικά. Οι μεταλλάξεις είναι 1010G>A και 1024C>T. Το 1010G ανήκει στην τριπλέτα 336 και βρίσκεται σε πλούσιο GC περιβάλλον (10 από τα 15 νουκλεοτίδια). Το 1024C ανήκει στην τριπλέτα 342 έχει 8 από τα 15 νουκλεοτίδια GC. Επίσης ανήκει σε θέση CpG. Σε αντίθεση με την τριπλέτα 267, η τριπλέτα 342 περιέχει CpG και είναι θέση αυξημένης μετάλλαξης για τα σωματικά κύτταρα και όχι τα γενετικά. Πάντως αυτό το σημείο χρήζει περαιτέρω διερεύνηση από την επιστημονική κοινότητα.

2. Στα παρακάτω σχήματα γίνεται σύγκριση μεταξύ γενετικών και σωματικών μεταλλάξεων ως προς συγκεκριμένες μεταβλητές, με τη χρήση όλων των εγγραφών των γενετικών μεταλλάξεων (1424) και όλων των σωματικών (21587).

(α) Γενετικές μεταλλάξεις



(β) Σωματικές μεταλλάξεις



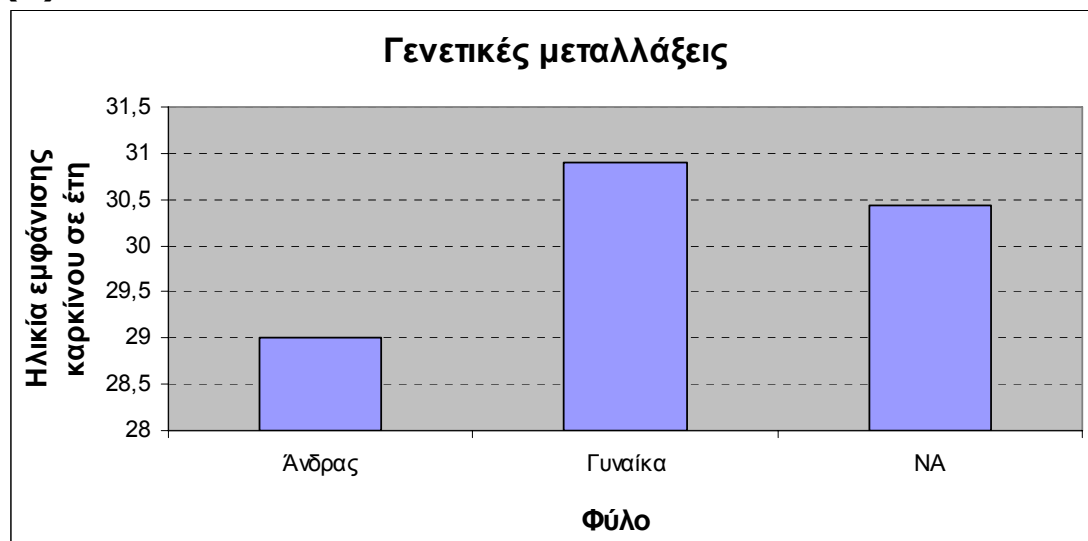
**Σχήμα 8. Κατηγορίες γενετικών και σωματικών μεταλλάξεων στο γονίδιο p53.**



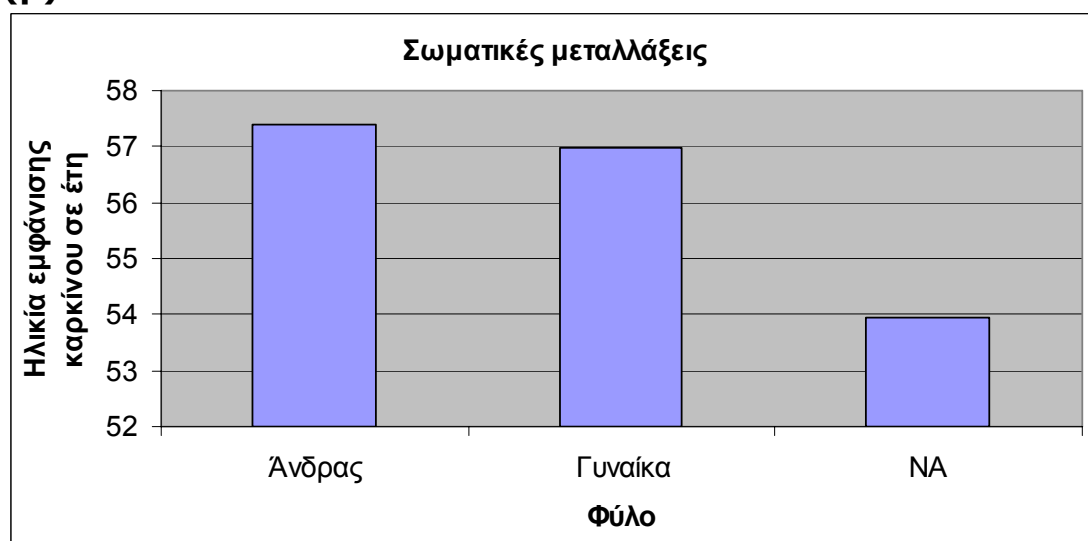
Όπως αναμένεται, οι πιο συχνές είναι οι παρανοηματικές (missense) μεταλλάξεις, που αποτελούν τα 3/4 των σωματικών και τα 2/3 των γενετικών μεταλλάξεων. Από τη βιβλιογραφία που μελετήθηκε οι 37 από τις 40 σιωπηρές μεταλλάξεις της γενετικής βάσης δεδομένων ανήκουν τελικά στην κατηγορία ματίσματος (splice). [116] [117] [118] Επομένως στα γενετικά κύτταρα υπάρχουν μόνο 3 μεταλλάξεις σιωπηρές και οι υπόλοιπες ανήκουν στην κατηγορία ματίσματος. Αυτό σημαίνει ότι οι γενετικές μεταλλάξεις έχουν ποσοστό σχεδόν 0% σιωπηρών μεταλλάξεων σε αντίθεση με τις σωματικές μεταλλάξεις που έχουν ποσοστό 4%. Επίσης οι γενετικές μεταλλάξεις έχουν τελικά ποσοστό 8% ματίσματος ενώ οι σωματικές μεταλλάξεις έχουν ποσοστό 2%.

### Συσχέτιση φύλου και μεταλλάξεων

(α)



(β)

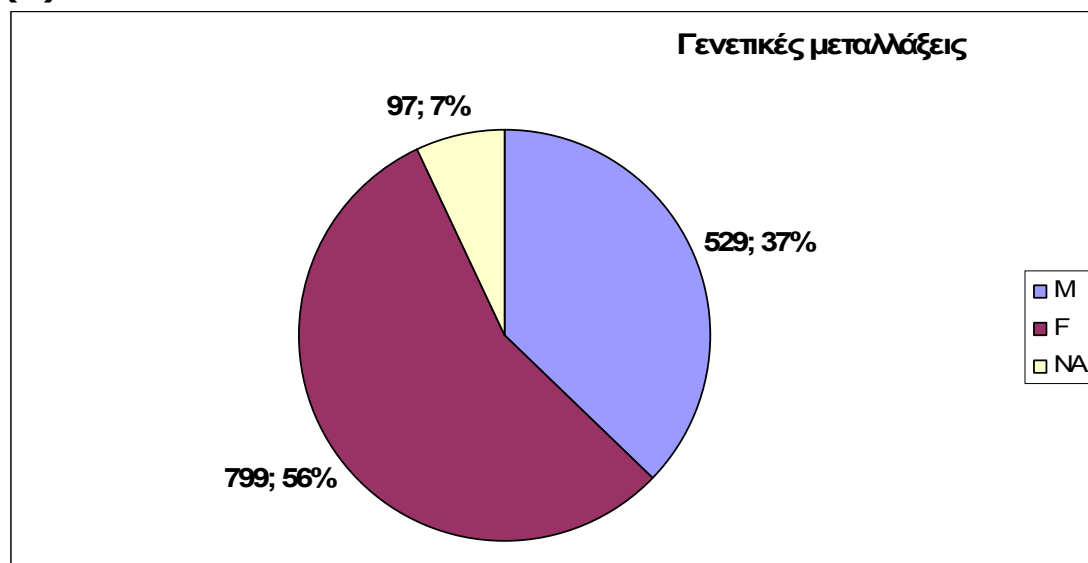


**Σχήμα 9. Συσχέτιση φύλου με ηλικία εμφάνισης καρκίνου στα γενετικά και στα σωματικά κύτταρα**

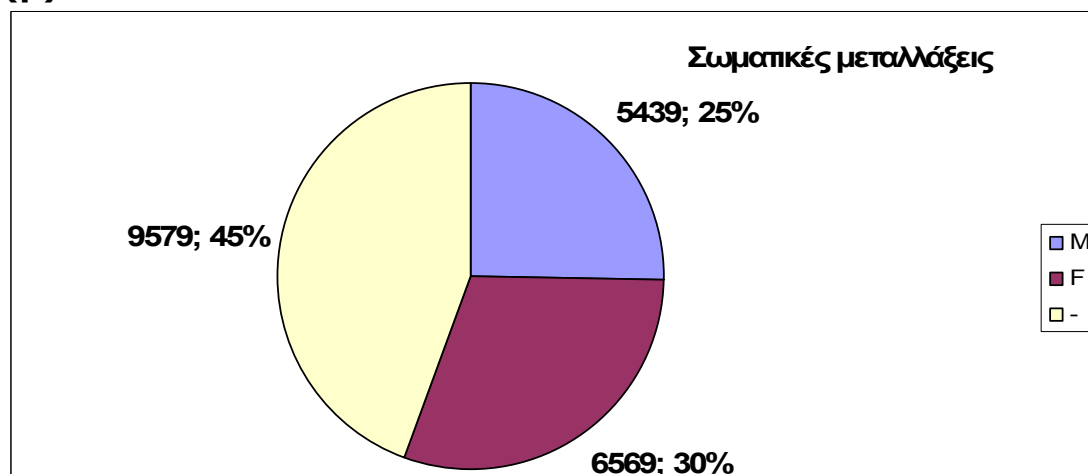
Παρατηρούμε ότι στις γενετικές μεταλλάξεις τα αρρενα άτομα έχουν προδιάθεση να αναπτύξουν καρκίνο 2 έτη νωρίτερα σε σύγκριση με τα θήλεα. Αντίθετα, στις σωματικές μεταλλάξεις φαίνεται ότι είναι παρόμοια η μέση ηλικία εμφάνισης.

Η στήλη NA εκπροσωπεί άτομα των οποίων δε γνωρίζουμε το φύλο. Στις γενετικές μεταλλάξεις βρίσκεται ανάμεσα στο μέσο όρο ανδρών και γυναικών, ενώ στις σωματικές βρίσκεται χαμηλότερα από το μέσο όρο και των δύο. Αυτό σημαίνει ότι πιθανότατα αν γνωρίζαμε το φύλο των NA και οι δύο μέσοι όροι στα σωματικά να ήταν χαμηλότεροι.

(α)



(β)



**Σχήμα 10. Σύγκριση ποσοστού των δύο φύλων στις γενετικές και σωματικές μεταλλάξεις.**

M=Άρρεν φύλο, F=Θήλυ φύλο, NA και - =Δεν είναι γνωστό το φύλο

Οι εγγραφές των γενετικών μεταλλάξεων είναι 1.424 και από αυτές μόνο 54 αρρενες και 41 θήλεις δεν εμφάνισαν ακόμη νεόπλασμα. Οι εγγραφές των σωματικών μεταλλάξεων είναι 21.587 και όλοι έχουν εμφανίσει νεόπλασμα.

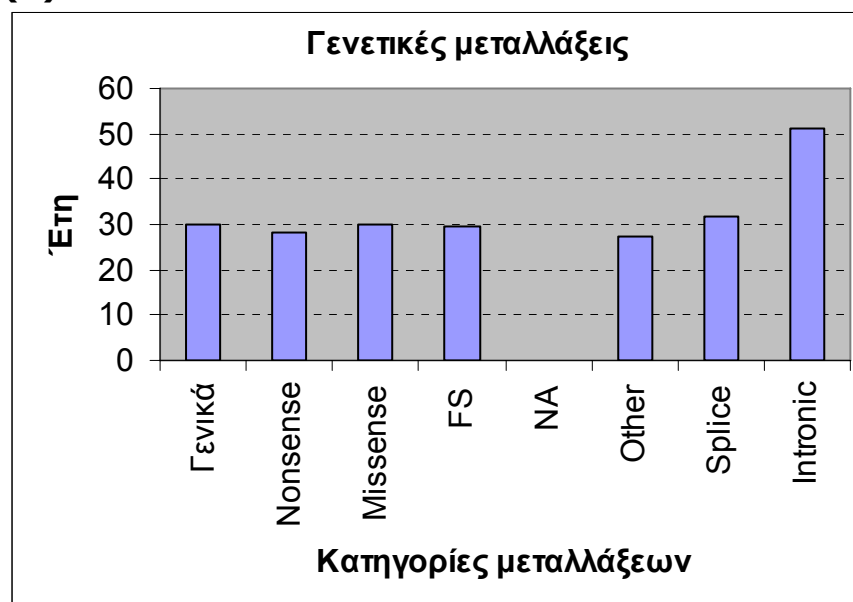
Παρατηρούμε ότι στα γενετικά κύτταρα γνωρίζουμε σε ποσοστό 97% το φύλο και υπερτερούν στις εγγραφές οι γυναίκες. Στα σωματικά κύτταρα γνωρίζουμε το φύλο μόνο σε ποσοστό 55% και οι εγγραφές μεταξύ των φύλων είναι περίπου ίσες.

Αν αφαιρέσουμε τους φορείς χωρίς εκδήλωση νεοπλασματος και εφαρμόσουμε τη δοκιμασία  $\chi^2$  μεταξύ των γενετικών και σωματικών καταγραφών των δύο φύλων έχουμε στατιστικώς σημαντική διαφορά με  $p = 0,0000$ . Αυτό σημαίνει ότι επηρεάζονται διαφορετικά τα δύο φύλα, αν η μετάλλαξη είναι γενετική ή σωματική.

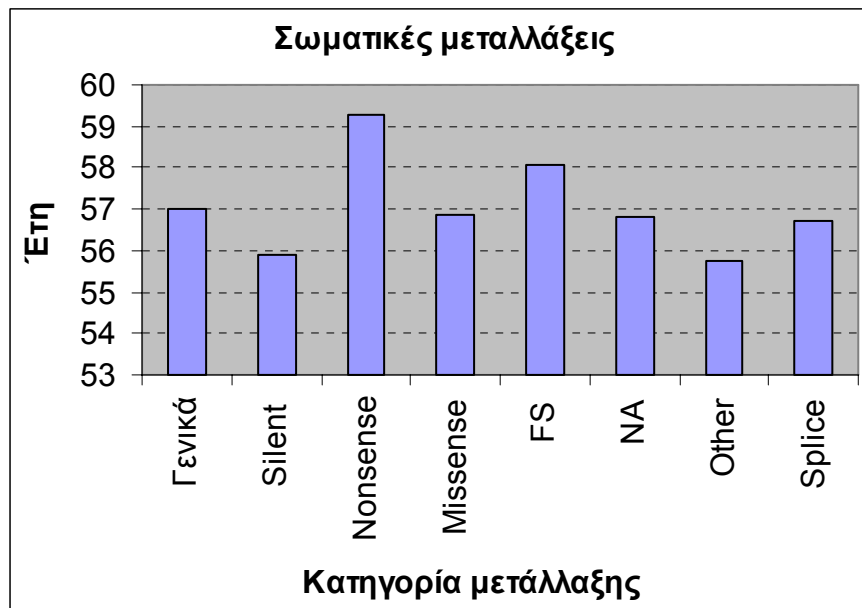
Στις γενετικές μεταλλάξεις υπερτερούν οι θήλεις. Άρα είτε οι θήλεις επιζούν μετά τη γέννηση λόγω του δεύτερου X χρωμοσώματος έναντι των αρρένων και είναι περισσότερες ως φορείς άρα και ως νοσούντες, είτε οι αρρενες λόγω κάποιου άγνωστου χαρακτηριστικού είναι πιο ανθεκτικοί στην εμφάνιση νεοπλασματος από γενετική μετάλλαξη.

### Μέσος όρος ηλικίας εμφάνισης καρκίνου

(α)



(β)



**Σχήμα 11. Σύγκριση μέσου όρου ηλικίας εμφάνισης νεοπλασμάτων στις γενετικές και σωματικές μεταλλάξεις ανάλογα με την κατηγορία μετάλλαξης.**

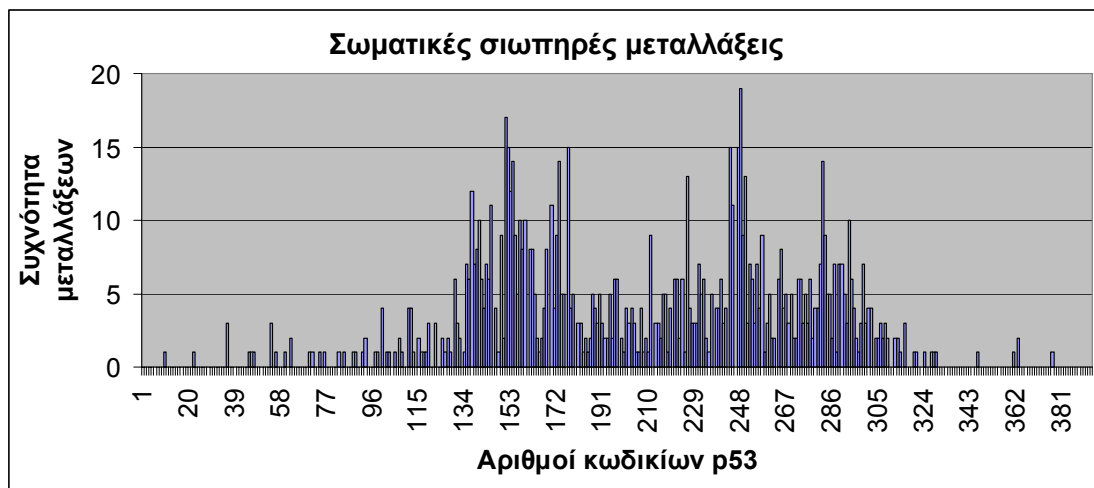
Όπως αναφέρθηκε νωρίτερα, η ηλικία εμφάνισης καρκίνου είναι γύρω στα 30 έτη για τα άτομα με γενετικές μεταλλάξεις και γύρω στα 55 έτη για άτομα με σωματικές μεταλλάξεις. Οι γενετικές μεταλλάξεις του *p53* είναι εξαρχής συσσωρευμένες σε όλα τα κύτταρα και εμφανίζουν τα νεοπλάσματα 30 έτη αργότερα. Επομένως, πιθανότατα, οι σωματικές μεταλλάξεις συσσωρεύονται στον ίδιο βαθμό με τις γενετικές στα κύτταρα στα 25 έτη ζωής και προκαλούν νεοπλάσματα 30 έτη αργότερα. Όταν ένα άτομο είναι φορέας γενετικής μετάλλαξης του *p53* γονιδίου, θα εμφανίσει καρκίνο κατά μέσο όρο 20 έτη νωρίτερα από άλλο άτομο που θα υποστεί σωματική μετάλλαξη του *p53*.

Ο μέσος όρος εμφάνισης καρκίνου στο γενικό πληθυσμό είναι τα 67 έτη. [132] Αυτό σημαίνει ότι οι φορείς των μεταλλάξεων του *p53* έχουν στατιστικώς σημαντική πτώση του μέσου όρου ηλικίας εμφάνισης καρκίνου και στις δύο σειρές κυττάρων.

Οι γενετικές μεταλλάξεις των ενδονίων (intronic) φαίνεται ότι επηρεάζουν σε μικρό βαθμό την παραγόμενη πρωτεΐνη και έχουν μέσο όρο εμφάνισης καρκίνου που προσεγγίζει τα σωματικά κύτταρα. Όμως πάλι είναι χαμηλός σε σχέση με το γενικό μέσο όρο που είναι τα 67 έτη. [132] Επομένως και οι μεταλλάξεις των ενδονίων επηρεάζουν την επιβίωση των φορέων.

Οι σιωπηρές μεταλλάξεις έχουν τον ίδιο μέσο όρο ηλικίας με τα υπόλοιπα είδη στις σωματικές μεταλλάξεις. Λογικά αυτό δεν πρέπει

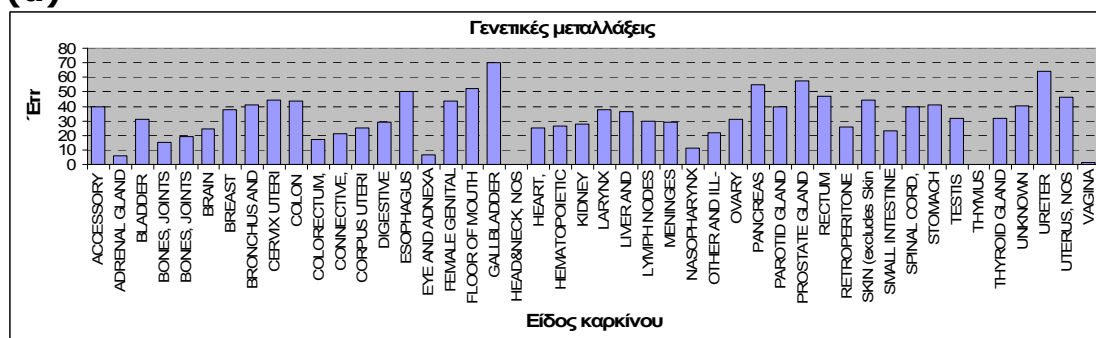
να συμβαίνει αφού οι σιωπηρές μεταλλάξεις διατηρούν αναλλοίωτη τη δομή της πρωτεΐνης. Μπορούμε να υποθέσουμε είτε ότι οι σιωπηρές μεταλλάξεις συνδυάζονται με άλλες μεταλλάξεις στο ίδιο γονίδιο, είτε ότι προκαλούν τη δημιουργία γενετικά ευάλωτων θέσεων που ξαναμεταλλάσσονται κατά τη διάρκεια της ζωής. Στα γενετικά κύτταρα βρήκαμε ότι οι σιωπηρές μεταλλάξεις συνδυάζονται με προσθήκη στο μεταγράφημα του ενδονίου 4 και επομένως δεν μπορούν να χαρακτηριστούν ως σιωπηρές.



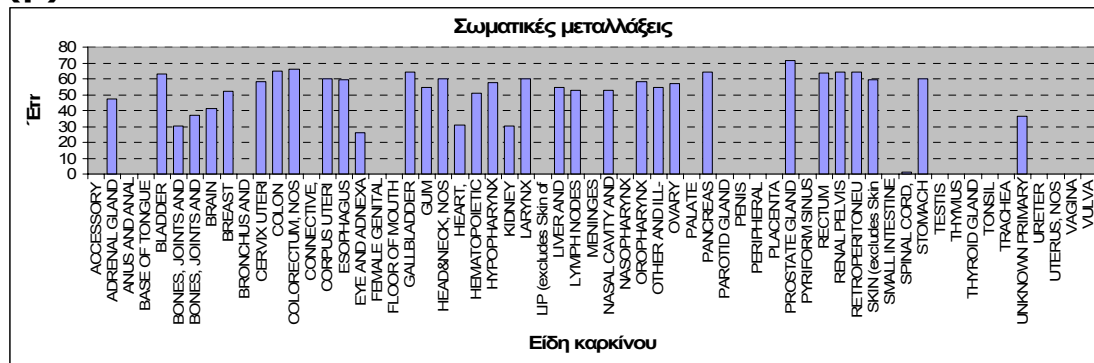
**Σχήμα 12. Καταγραφές σιωπηρών σωματικών μεταλλάξεων στην αλληλουχία του γονιδίου.**

Όπως γίνεται αντιληπτό, οι σιωπηρές μεταλλάξεις κατανέμονται σε όλη την έκταση του γονιδίου, κυρίως στα εξώνια 5-8, και αριθμούν συνολικά 951 εγγραφές. Αποτελούν δείκτη γενικευμένης γενετικής αστάθειας των σωματικών κυττάρων.

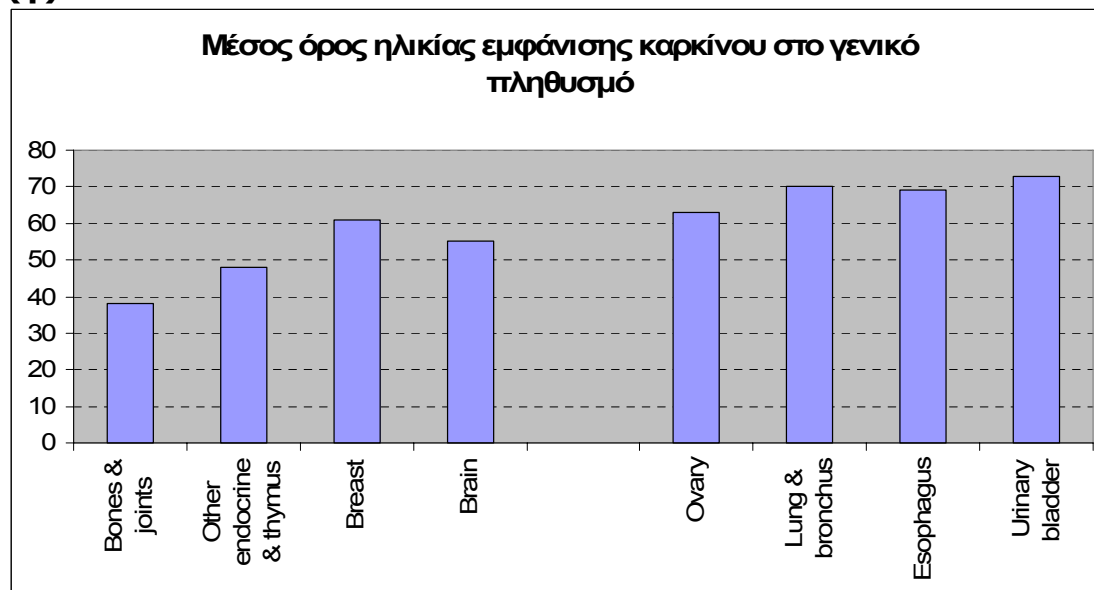
**(a)**



(β)



(γ)



**Σχήμα 13. Μέσος όρος ηλικίας εμφάνισης καρκίνου στις γεννητικές (α), στις σωματικές μεταλλάξεις (β) και στο γενικό πληθυσμό (γ).**

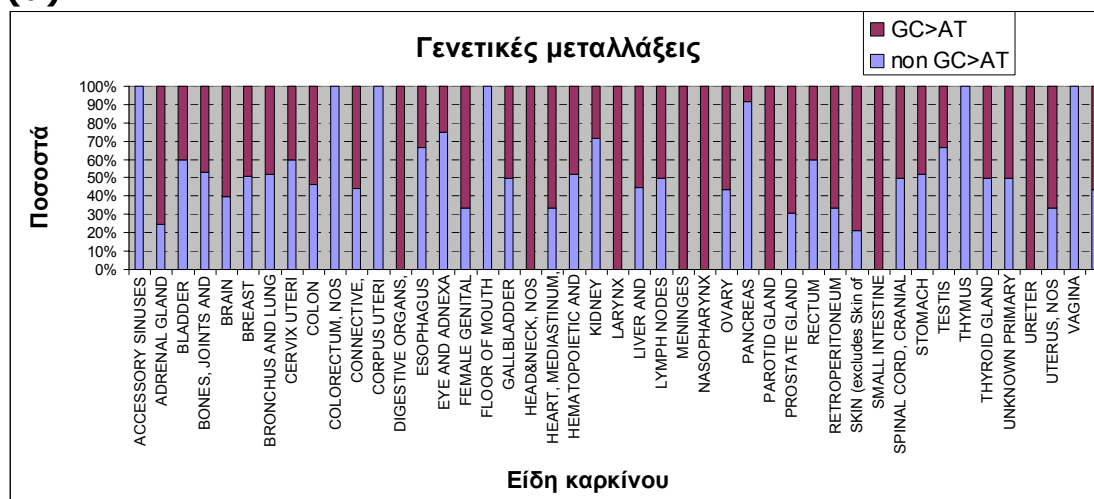
Ο μέσος όρος ηλικίας εμφάνισης καρκίνου στις γεννητικές μεταλλάξεις είναι 33,7 έτη. Ο μέσος όρος ηλικίας εμφάνισης καρκίνου στις σωματικές μεταλλάξεις είναι 52,2 έτη. Δυστυχώς δεν υπάρχουν ηλικιακά στοιχεία για όλα τα είδη του καρκίνου. Ο μέσος όρος εμφάνισης καρκίνου στο γενικό πληθυσμό είναι 67 έτη. [132]

Στον παρακάτω πίνακα απεικονίζονται τα νεοπλάσματα με τη μεγαλύτερη συχνότητα και ο μέσος όρος εμφάνισης καρκίνου.

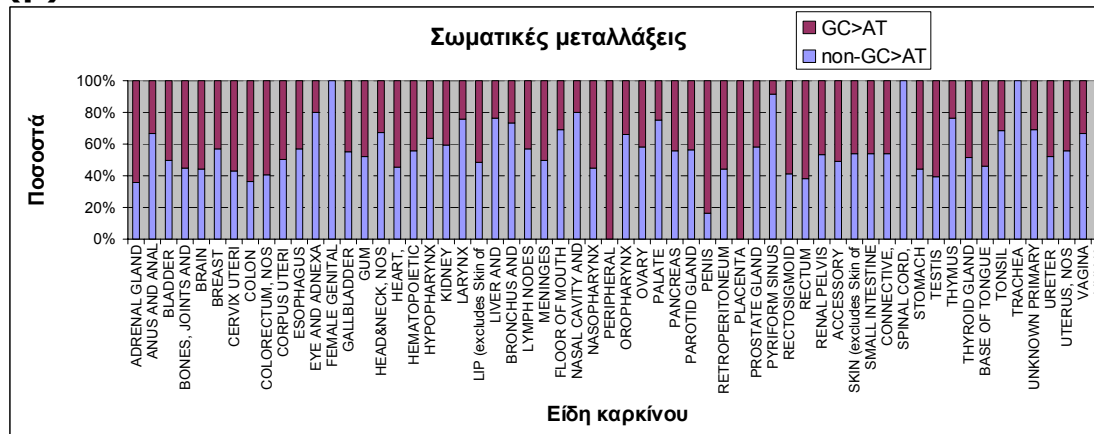
	Γενικός πληθυσμός μέσος όρος ηλικίας εμφάνισης καρκίνου		Γενετικά κύτταρα μέσος όρος ηλικίας εμφάνισης καρκίνου	Σωματικά κύτταρα μέσος όρος ηλικίας εμφάνισης καρκίνου
Other endocrine & thymus	48	ADRENAL GLAND	6,27	47,2
Urinary bladder	73	BLADDER	31,25	63,31
Bones & joints	38	BONES, JOINTS LIMBS, OTHER	17,4	33,09
Brain	55	BRAIN	24,65	41,33
Breast	61	BREAST	37,66	52,14
Lung & bronchus	70	BRONCHUS AND LUNG	40,97	62,61
Esophagus	69	ESOPHAGUS	50	59,1
		HEMATOPOIETIC	26,25	50,75
Ovary	63	Ovary	31,38	57,26

### Είδη καρκίνου στις γενετικές και σωματικές μεταλλάξεις

(α)



(β)



**Σχήμα 14. Συσχέτιση είδους καρκίνου και μεταλλάξεων που οφείλονται σε μεθυλίωση στα γενετικά και σωματικά κύτταρα.**

Είναι γνωστό ότι οι G:C>A:T μεταλλάξεις οφείλονται σε μεθυλίωση. Σε αυτές έχουν υπολογιστεί και οι μεταλλάξεις σε θέσεις CpG. Συμπερασματικά, εντοπίζονται διαφορές στο ποσοστό των μεταλλάξεων G:C>A:T μεταξύ σωματικών και γενετικών μεταλλάξεων με έμφαση στον κόλπο, θύμο, πάγκρεας, ορθό, σώμα μήτρας και γυναικεία γεννητικά όργανα. Από την άλλη υπάρχουν ομοιότητες στα ποσοστά του εγκεφάλου, του μαστού, του νεφρού και της κύστεως.

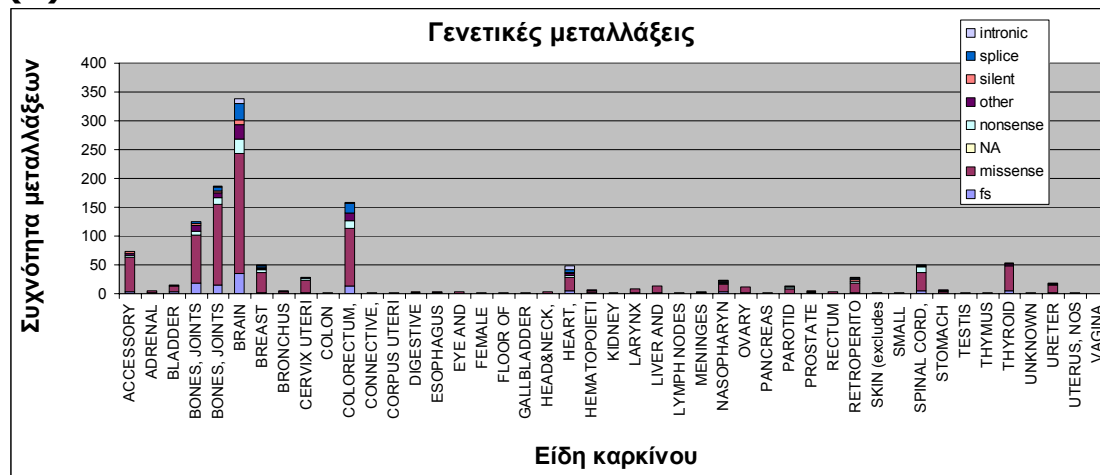
Στον παρακάτω πίνακα γίνεται η σύγκριση του βαθμού μεθυλίωσης για τα νεοπλάσματα με τις περισσότερες καταγραφές.

Topography	ΓΕΝΕΤΙΚΕΣ		Topography	ΣΩΜΑΤΙΚΕΣ		ΔΟΚΙΜΑΣΙΑ X2
	non G:C>A:T	G:C>A:T		non-G:C>A:T	G:C>A:T	
ADRENAL GLAND	18	56	ADRENAL GLAND	25	45	0,1355
BLADDER	3	2	BLADDER	509	521	0,6368
BONES, JOINTS	8	7	BONES, JOINTS	22	27	0,5667
BRAIN	74	113	BRAIN	637	799	0,2146
BREAST	172	167	BREAST	1265	944	0,024
BRONCHUS AND LUNG	26	24	BRONCHUS AND LUNG	1739	633	0,0008
CONNECTIVE, SUBCUTANEOUS, SOFT TISSUES	70	88	CONNECTIVE, SUBCUTANEOUS, SOFT TISSUES	153	131	0,0538
ESOPHAGUS	2	1	ESOPHAGUS	940	713	0,732
HEMATOPOIETIC	25	23	HEMATOPOIETIC	524	416	0,6185
OVARY	10	13	OVARY	717	512	0,1524
SKIN	6	22	SKIN	512	438	0,0007

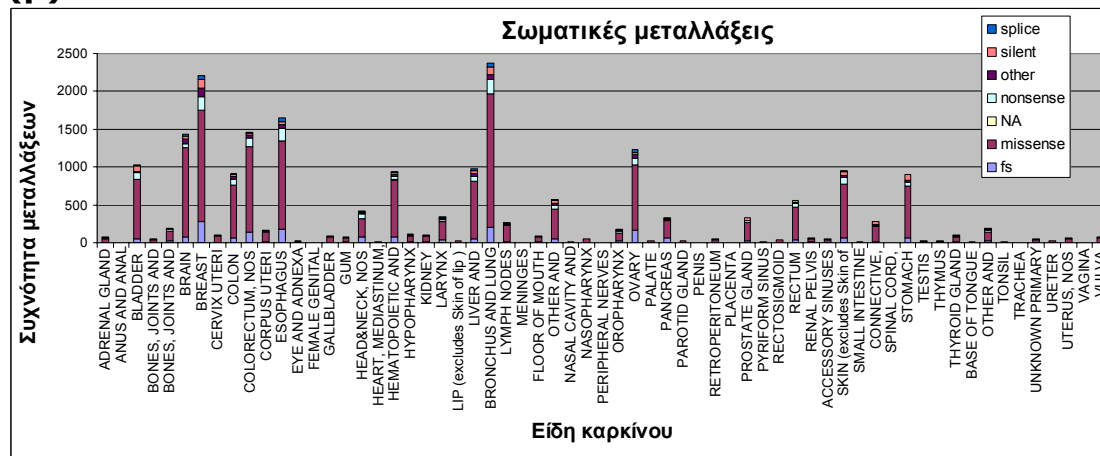
Παρατηρούμε ότι οι μεθυλιώσεις που συμβαίνουν μεταξύ των γενετικών και σωματικών κυττάρων δεν διαφέρουν στατιστικώς σημαντικά μεταξύ τους, με εξαίρεση τον καρκίνο του πνεύμονα, του μαστού και του δέρματος. Φαίνεται ότι τα καρκινογόνα προϊόντα του καπνού είναι εξαιρετικοί μεθυλιωτικοί παράγοντες και προσθέτουν τόσες πολλές μεταλλάξεις μεθυλίωσης στα σωματικά κύτταρα που διαφοροποιούν τελείως το πρότυπο μεθυλίωσης από αυτό των γεννητικών κυττάρων. Επίσης η ηλιακή ακτινοβολία και επαφή με χημικές ουσίες μπορεί να επιτείνουν την μεταλλαξιγένεση στον καρκίνο του δέρματος. Ο μαστός και ο συνδετικός ιστός παράγουν διαφορετικές ισομορφές της πρωτεΐνης TP53 και όπως φαίνεται οι ισομορφές αυτές είναι πιο ευάλωτες σε επιγενετικές τροποποιήσεις.



(α)



(β)



### Σχήμα 15. Συσχέτιση είδους καρκίνου με το είδος μετάλλαξης στις γενετικές και σωματικές μεταλλάξεις.

Η πρώτη στήλη BONES,JOINTS αναφέρεται μόνο στα άκρα ενώ η δεύτερη στήλη αναφέρεται σε όλα τα υπόλοιπα σημεία του σώματος. Στη στήλη SKIN εξαιρείται το δέρμα του όσχεου. Η στήλη OTHER που βρίσκεται δίπλα στη στήλη FLOOR OF MOUTH αναφέρεται στα υπόλοιπα σημεία της στοματικής κοιλότητας. Η στήλη OTHER που βρίσκεται δίπλα στη στήλη BASE OF TONGUE αναφέρεται στα υπόλοιπα σημεία της γλώσσας.

Παρατηρούμε ότι η παρουσία κανενός τύπου μετάλλαξης δε σχετίζεται αποκλειστικά με την εμφάνιση συγκεκριμένου νεοπλασματος. Επίσης παρατηρούμε ότι τα νεοπλασματα που σχετίζονται με περιβαλλοντικές επιδράσεις έχουν περισσότερες σωματικές καταγραφές.

Παρατηρούμε μεγάλες διαφορές ιδιαίτερα στο καρκίνο του πνεύμονα που είναι αυξημένος στα σωματικά κύτταρα. Πρέπει να ληφθεί υπόψη ότι οι δύο πληθυσμοί δεν έχουν το ίδιο μέγεθος, επειδή οι καταγραφές των γενετικών μεταλλάξεων είναι αρκετά λιγότερες.

Στον παρακάτω πίνακα συγκρίνονται με τη βοήθεια της δοκιμασίας  $\chi^2$  οι συχνότητες των νεοπλασμάτων που έχουν τις περισσότερες καταγραφές στις γενετικές ή στις σωματικές μεταλλάξεις. Οι τιμές του γενικού πληθυσμού λήφθηκαν από το **Cancer Statistics Review 1975-2002, National Cancer Institute, USA**. [132]

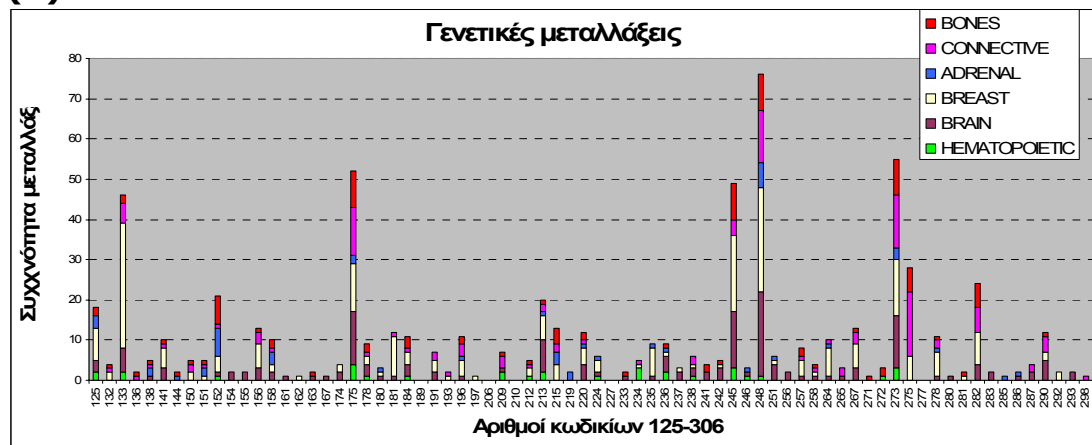
Γενικός πληθυσμός	Συχνότητα ανά 100.000 κατοίκους		Γενετικά κύτταρα (1424)	Δοκιμασία $\chi^2$ σε σχέση με το γενικό πληθυσμό	Σωματικά κύτταρα (21.587)	Δοκιμασία $\chi^2$ σε σχέση με το γενικό πληθυσμό
Other endocrine & thymus	0,7	ADRENAL GLAND	74	0,0000	25	0,0000
Urinary bladder	20,3	BLADDER	5	0,0000	509	0,0000
Bones & joints	0,9	BONES, JOINTS AND ARTICULAR CARTILAGE OF LIMBS	15	0,0000	22	0,0000
		BONES, JOINTS AND ARTICULAR CARTILAGE OF OTHER AND UNSPECIFIED SITES	125	0,0000	94	0,0000
Brain	6	BRAIN	187	0,0000	638	0,0000
Breast	73,3	BREAST	339	0,0000	1265	0,0000
Lung & bronchus	61	BRONCHUS AND LUNG	50	0,0000	1739	0,0000
Esophagus	4,5	ESOPHAGUS	3	0,0000	940	0,0000
Breast ( <i>in situ</i> )	16,6	HEMATOPOIETIC AND RETICULOENDOTHELIAL SYSTEMS	48	0,0000	524	0,0000
Ovary	7,6	OVARY	23	0,0000	717	0,0000

Προκύπτει ότι οι φορείς σωματικών και γενετικών μεταλλάξεων έχουν στατιστικώς πολύ σημαντική αύξηση της συχνότητας των νεοπλασμάτων σε σύγκριση με το γενικό πληθυσμό για όλα τα είδη καρκίνου.

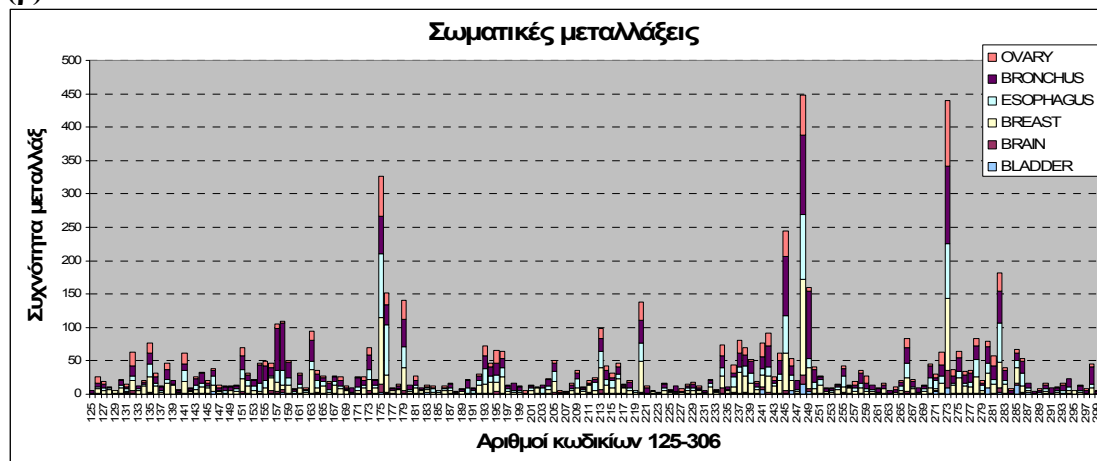
	Γενετικά κύτταρα	Σωματικά κύτταρα	Δοκιμασία $\chi^2$ μεταξύ γενετικών και σωματικών
	Καταγραφές μεταλλάξεων	Καταγραφές μεταλλάξεων	
ADRENAL GLAND	74	25	0,0000
BLADDER	5	509	0,0000
BONES, JOINTS LIMBS	15	22	0,0000
BONES, JOINTS OTHER	125	94	0,0000
BRAIN	187	638	0,0000
BREAST	339	1265	0,0000
BRONCHUS AND LUNG	50	1739	0,0000
ESOPHAGUS	3	940	0,0000
HEMATOPOIETIC	48	524	0,0318
OVARY	23	717	0,0006

Από τη σύγκριση μεταξύ γενετικών και σωματικών μεταλλάξεων προκύπτει ότι διαφέρουν οι συχνότητες των νεοπλασμάτων στατιστικώς πολύ σημαντικά. Αυτό σημαίνει ότι όποιος είναι φορέας γενετικών μεταλλάξεων έχει προδιάθεση για ανάπτυξη διαφορετικών νεοπλασμάτων σε σχέση με τις σωματικές.

(α)



(β)



**Σχήμα 16. Μεταλλάξεις κωδικίων στην αλληλουχία του p53 ανάλογα με το είδος καρκίνου, στα γενετικά και σωματικά κύτταρα.**

Μελετούμε τα είδη καρκίνου με τις περισσότερες συνολικές καταγραφές στις γενετικές ή σωματικές βάσεις μεταλλάξεων.

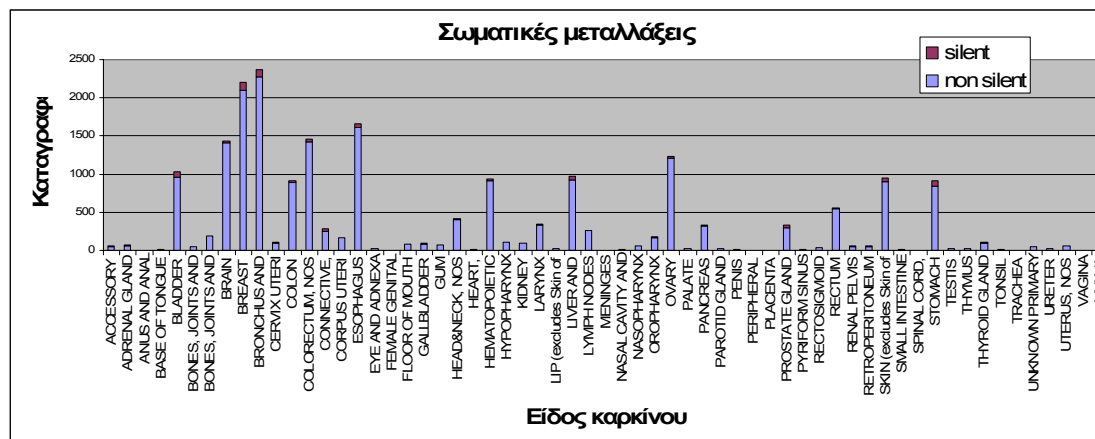
Γενετικά είδη καρκίνου	Σωματικά είδη καρκίνου
<ul style="list-style-type: none"> <li>✓ Bones</li> <li>✓ Connective</li> <li>✓ Adrenal</li> <li>✓ Breast</li> <li>✓ Brain</li> <li>✓ Hematopoietic</li> </ul>	<ul style="list-style-type: none"> <li>✓ Ovary</li> <li>✓ Bronchus</li> <li>✓ Esophagus</li> <li>✓ Breast</li> <li>✓ Brain</li> <li>✓ Bladder</li> </ul>

Παρατηρούμε ότι οι γενετικές και σωματικές μεταλλάξεις έχουν από κοινού πολλές καταγραφές μόνο στους καρκίνους του εγκεφάλου και του μαστού. Η απουσία καταγραφής μεταλλάξεων καρκίνου σε όλα τα κωδικία είναι μία ένδειξη ότι η γενετική βάση δεδομένων είναι ημιτελής ή ότι επιδρούν διαφορετικοί μηχανισμοί γενετικής αστάθειας.

Τα κωδικία που φαίνεται ότι μεταλλάσσονται περισσότερο σε σωματικούς καρκίνους είναι:

Καρκίνος	Κωδικία
Πνεύμονας	151-159, 175, 179, 245, 248, 249, 273, 282
Οισοφάγος	175, 176, 245, 248, 273, 282
Ουροδόχος κύστη	248, 273, 280, 285

Επίσης, παρατηρούμε στις γενετικές μεταλλάξεις ότι υπάρχουν αυξημένες πιθανότητες για εμφάνιση καρκίνου του μαστού αν συμβεί μετάλλαξη στο κωδικό 133 (60%), 175 (20%), 181 (80%), 245 (40%), 248 (40%), 264 (70%) και 273 (25%).



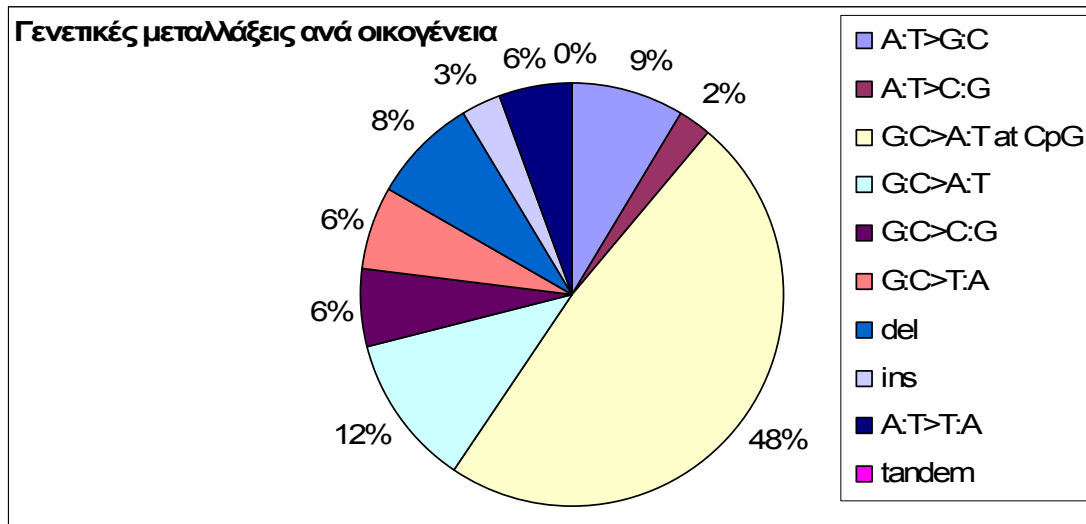
**Σχήμα 17. Καταγραφές σιωπηρών και υπόλοιπων μεταλλάξεων σε σχέση με το είδος του καρκίνου στα σωματικά κύτταρα.**

Σωματικά κύτταρα		
Τοπογραφία νεοπλασίας με καταγραφές πάνω από 800	με συνολικές	Ποσοστό σιωπηρών μεταλλάξεων
BRAIN		2%
BREAST		5%
COLON		3%
COLORECTUM, NOS		2%
ESOPHAGUS		2%
HEMATOPOIETIC AND RETICULOENDOTHELIAL SYSTEMS		3%
LIVER AND INTRAHEPATIC BILE DUCTS		5%
BRONCHUS AND LUNG		4%
OVARY		2%
SKIN (excludes Skin of vulva, Skin of penis, Skin of scrotum )		6%
STOMACH		8%

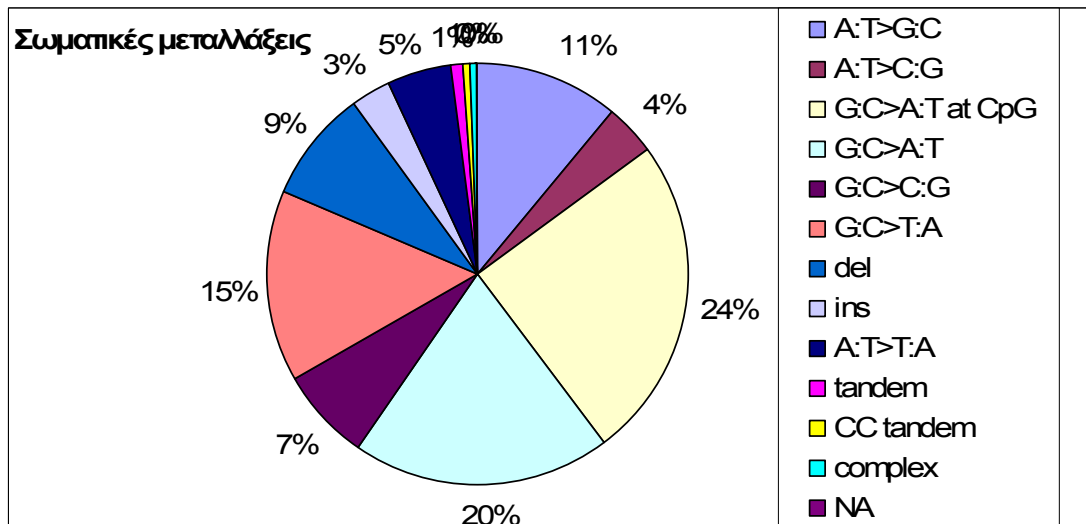
Παρατηρούμε ότι οι καταγραφές των σιωπηρών μεταλλάξεων διαφέρουν σημαντικά από τις αναμενόμενες που είναι κάτω του 1%.

**3. Για τη δημιουργία των επόμενων σχημάτων χρησιμοποιήθηκαν όλες οι καταγραφές των σωματικών μεταλλάξεων και μία καταγραφή ανά οικογένεια των γενετικών μεταλλάξεων. Αυτό σημαίνει ότι κάθε οικογενειακή μετάλλαξη αντιπροσωπεύεται μία φορά. Όπως αναφέρθηκε στην μεθοδολογία της εργασίας, αυτό γίνεται για να μην εξαρτάται η συχνότητα κάθε μετάλλαξης από τον αριθμό των απογόνων των φερέων.**

**(α)**



**(β)**



**Σχήμα 18. Τύποι γενετικών και σωματικών μεταλλάξεων.**

Εφόσον υπάρχει μεθυλίωση στις θέσεις CpG, οι μεταλλάξεις G:C>A:T θα πρέπει να διακριθούν σε G:C>A:T at CpG και G:C>A:T non CpG. Οι σημαντικές διαφορές σημειώνονται με κόκκινο:

	Γεννητικά κύτταρα	Σωματικά κύτταρα
Τύπος μετάλλαξης	Τιμή	Τιμή
A:T>G:C	9%	11%
A:T>C:G	2%	4%
G:C>A:T at CpG	48%	25%
G:C>A:T	12%	20%
G:C>C:G	6%	7%
G:C>T:A	6%	15%
del	8%	9%
ins	3%	3%
A:T>T:A	6%	5%
tandem	0%	1%
CC tandem	0%	1%
complex	0%	0%
NA	0%	0%

Τύπος μετάλλαξης	Σωματικές μεταλλάξεις (21.587)	Γενετικές μεταλλάξεις (282)	Δοκιμασία $\chi^2$ , p=
G:C>A:T at CpG	5350	136	0,0000
G:C>A:T	4246	33	0,0045
G:C>T:A at CpG	500	3	0,1706
G:C>T:A	2637	15	0,0013
A:T>C:G	827	53	0,0000

Οι διαφορές αυτές είναι στατιστικώς πολύ σημαντικές, όπως φαίνεται με την εφαρμογή του κριτηρίου  $\chi^2$  στις απόλυτες τιμές των καταγραφών τους. Εκτός από τις μεταλλάξεις G:C>T:A επειδή είναι πολύ λίγες οι καταγραφές στα γενετικά κύτταρα. Άρα αυτό φανερώνει ότι οι δύο κυτταρικές σειρές μεταλλάσσονται με διαφορετικό τρόπο και πιθανότατα με την επίδραση διαφορετικών παραγόντων.

Κωδικίο	Μεταλλαξιμότητα	Κωδικίο	Μεταλλαξιμότητα	Κωδικίο	Μεταλλαξιμότητα	Κωδικίο	Μεταλλαξιμότητα
TTT	1,333333	TCT	0,285714	CAT	0,7	TGT	0,333333
TTC	– (0/0)	TCC	1,333333	CAC	0,166667	TGC	1,666667
TTA	– (0/0)	TCA	– (0/0)	TAA	– (0/0)	TGA	– (0/0)
TTG	0,7	TCG	– (0/0)	TAG	– (0/0)	TGG	0,2
CTT	– (0/0)	CCT	0,333333	TAT	0,368421	CGT	0,333333
CTC	– (0/0)	CCC	0,333333	TAC	0,090909	CGC	0,368421
CTA	0,833333	CCA	– (0/0)	CAA	0,428571	CGA	0,181818
CTG	0,2	CCG	1,333333	CAG	0,333333	CGG	1,333333
ATT	– (0/0)	ACT	– (0/0)	AAT	– (0/0)	AGT	0,307692
ATC	0,833333	ACC	0,583333	AAC	0,166667	AGC	0,2
ATA	– (0/0)	ACA	0,2	AAA	0,583333	AGA	0,333333
ATG	0,583333	ACG	0,307692	AAG	0,2	AGG	0,307692
GTT	– (0/0)	GCT	– (0/0)	GAT	0	GGT	0,2
GTC	0,333333	GCC	0,166667	GAC	0,2	GGC	0,368421
GTA	– (0/0)	GCA	0,333333	GAA	0,181818	GGA	0
GTG	0	GCG	– (0/0)	GAG	0,125	GGG	0,307692

**Σχήμα 19. Η σταθμισμένη μεταλλαξιμότητα των γεννητικών κυττάρων με μοναδικές εγγραφές ανά οικογένεια.**

Στον παραπάνω πίνακα υπολογίζεται η σταθμισμένη μεταλλαξιμότητα των κωδικίων του  $p53$  ως κλάσμα με αριθμητή τον αριθμό των μεταλλάξεων κάθε κωδικίου και παρονομαστή τον αριθμό εμφάνισης του κωδικίου στο γονίδιο.

Η συχνότητα του παραπάνω πίνακα επηρεάζεται από τη μεγάλη ή τη μικρή αξία (τοξικότητα) της μετάλλαξης. Επομένως, οι παραπάνω συχνότητες είναι αποτέλεσμα και της τοξικότητας της μετάλλαξης και της μεταλλαξιμότητας του κωδικίου.

Όπως φαίνεται υψηλή σταθμισμένη μεταλλαξιμότητα στις γενετικές μεταλλάξεις έχουν τα κωδικία CCG, CGG, TCC, TGC, TTT.

Η μεταλλαξιμότητα της TAC στα γενετικά κύτταρα είναι 0,090, η χαμηλότερη όλων των κωδικίων. Αυτό οφείλεται εν μέρει στο γεγονός ότι η μετάλλαξη μεθυλίωσης της TAC (C>T) οδηγεί στο κωδικίο TAT που κωδικοποιεί το ίδιο αμινοξύ, την τυροσίνη, και η μετάλλαξη είναι σιωπηρή.

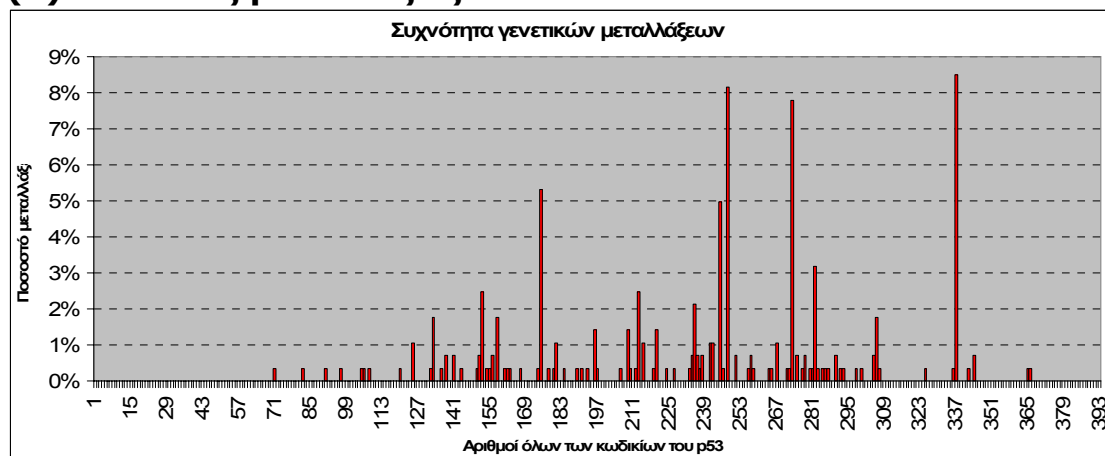
Επίσης, παρατηρούμε ότι τα δινουκλεοτίδια έχουν διαφορετική βιοχημική συμπεριφορά και διαφορετική τάση για μετάλλαξη ανάλογα με την τριπλέτα στην οποία συμμετέχουν. Επομένως πρέπει



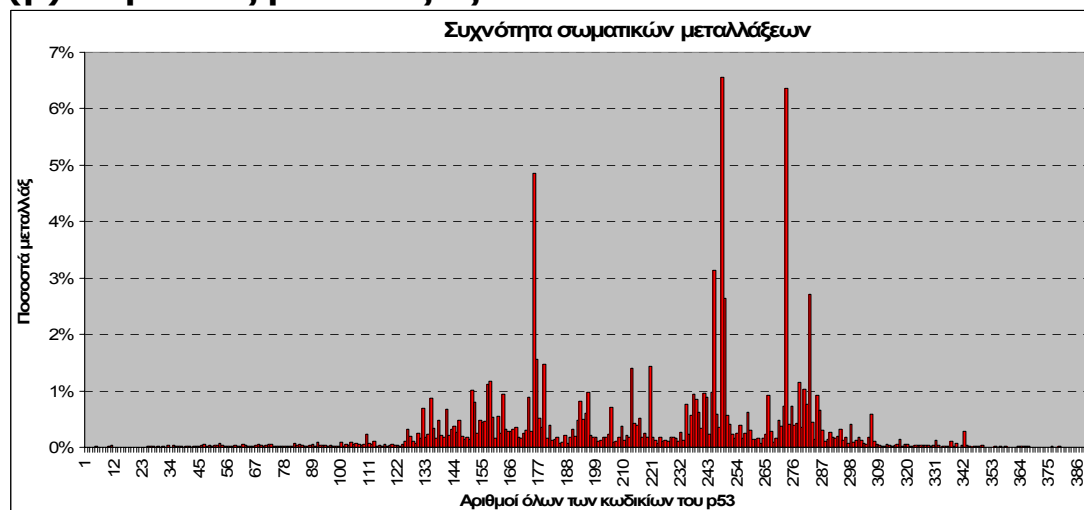
να μελετήσουμε πρότυπα νουκλεοτιδίων μεγαλύτερα από δύο για να εξάγουμε ασφαλή συμπεράσματα.  
Για παράδειγμα,

Κωδικία	Σταθμισμένη μεταλλαξιμότητα	Κωδικία	Σταθμισμένη μεταλλαξιμότητα
CCG	1,3	ACG	0,36
CGG	1,3	CGC	0,36
TCC	1,3	CCC	0,33
TGC	1,6	AGC	0,2
TTG	0,7	TTT	1,3

### (α) Γενετικές μεταλλάξεις



### (β) Σωματικές μεταλλάξεις



**Σχήμα 20. Κατανομή των μεταλλάξεων στην αλληλουχία του DNA στις γενετικές και σωματικές μεταλλάξεις.**

Η βάση δεδομένων των γενετικών μεταλλάξεων είναι κατά πολύ μικρότερη της βάσης των σωματικών. Έτσι τα γενετικά δεδομένα ανά οικογένεια περιέχουν 282 εγγραφές και τα σωματικά 21.587. Επομένως ορισμένα σαφώς καθοριζόμενα hotspots των σωματικών να μην είναι εμφανή στα γεννητικά λόγω έλλειψης εγγραφών και

ορισμένα hotspots των γεννητικών να εξαλειφθούν σιγά σιγά καθώς αναπτύσσεται η βάση των δεδομένων.

Από τη σύγκριση των δύο γραφημάτων φαίνεται ότι η επί τοις εκατό (%) συμμετοχή μεταλλάξεων σε ορισμένα κωδικία είναι παρόμοια στις βάσεις γενετικών και σωματικών μεταλλάξεων, ενώ σε άλλα διαφέρει σημαντικά.

Για την εύρεση των hotspots χρησιμοποιήθηκαν δύο μέθοδοι και έπειτα συγκρίθηκαν τα αποτελέσματα.

Στην πρώτη μέθοδο ορίστηκαν αυθαίρετα ως hotspots όσα κωδικία έχουν ποσοστό άνω του 1,2% του συνολικού αριθμού μεταλλάξεων. Μ' αυτό τον τρόπο στα σωματικά κύτταρα ξεχωρίζουν τα κωδικία: 175 (4,86%), 176 (1,55%), 179 (1,49%), 213 (1,41%), 220 (1,43%), 245 (3,15%), 248 (6,55%), 249 (2,7%), 273 (6,37%), 282 (2,71%).

Στα γενετικά κύτταρα προκύπτουν τα εξής hotspots: 133 (1,77%), 152 (2,48%), 158 (1,77%), 175 (5,32%), 196 (1,42%), 209 (1,42%), 213 (2,48%), 220 (1,42%), 235 (2,13%), 245 (4,96%), 248 (8,16%), 273 (7,8%), 282 (3,19%), 306 (1,77%), 337 (8,51%).

Η δεύτερη μέθοδος υπολογισμού των hotspots πραγματοποιήθηκε με τη χρήση εξειδικευμένων μαθηματικών εξισώσεων. Ο υπολογισμός έγινε με τη βοήθεια του προγράμματος iMARS που χρησιμοποιεί αυτές τις εξισώσεις. Λόγω των περιορισμών του προγράμματος αυτές είναι δυνατόν να χρησιμοποιηθούν μόνο στο φάσμα των γενετικών μεταλλάξεων. Τα αποτελέσματα που προκύπτουν είναι: η γενετική θέση 454 (C) που εντοπίζεται στην τριπλέτα 152, 637(C) που εμπεριέχεται στην τριπλέτα 213, η θέση 703 (A) που βρίσκεται στην τριπλέτα 235, η θέση 742 (C) που εμπεριέχεται στην 248 τριπλέτα, οι θέσεις 817 (C) και 818 (G) που εντοπίζονται στην 273 τριπλέτα και η θέση 844 (C) που βρίσκεται στην 282 τριπλέτα.

### Hotspots γενετικών μεταλλάξεων

Iterations = 10000

Αριθμός νουκλεοτιδίου	Τιμή P
454	0,0164
637	0,0134
703	0,0405
742	0,0001
817	0,0001
844	0,0018

Παρατηρούμε ότι η στατιστική συμφωνεί πλήρως με τις προηγούμενες μεθόδους, δίνει όμως λιγότερα hotspots ως στατιστικώς σημαντικά.

### Ταξινόμηση hotspot μεταλλάξεων

Με κόκκινο συμβολίζονται τα hotspots που δεν έχουν αντίστοιχο στην άλλη σειρά:

**Σωματικά:** 175, 176, 179, 213, 220, 245, 248, 249, 273 και 282.

**Γενετικά:** 133, 152, 158, 175, 196, 209, 213, 220, 235, 245, 248, 273, 282, 306, 337.

**iMARS για γενετικά:** 152, 213, 235, 248, 273, 282.

Σημαντικός αριθμός αντίστοιχων κωδικίων φαίνεται να μεταλλάσσεται με την ίδια συχνότητα και σε γενετικά και σε σωματικά κύτταρα. Αυτό γίνεται εμφανές στα hotspots, όπου υπάρχει μεγάλη ομοιότητα στις θέσεις 175, 245, 248, 273 και 282. Εξαιρέση αποτελούν οι θέσεις 249 στα σωματικά, 133, 275 και 337 στα γενετικά. Σ' αυτές τα hotspots δεν έχουν αντίστοιχα στην άλλη σειρά. Τα hotspot ξεχωρίζουν χάρη στη συνέργεια δύο χαρακτηριστικών: της υψηλής μεταλλαξιμότητας και της μεγάλης τοξικότητας της μετάλλαξης. Αντίθετα, στα κωδικία που δεν είναι hotspots, είναι πιθανό είτε να μη συμβαίνουν μεταλλάξεις σ' αυτές τις θέσεις είτε να συμβαίνουν μεταλλάξεις αλλά να μην προκαλούν καρκίνο.

Παρατηρείται ότι οι γενετικές θέσεις των κωδικίων 152 και 158, που συνδέονται με μεγάλη τοξικότητα, είναι πολύ σταθερές (μικρός αριθμός καταγεγραμμένων μεταλλάξεων) και στα γενετικά και στα σωματικά κύτταρα. Αυτό ίσως προκύπτει λόγω επιδιόρθωσης της περιοχής από κατάλληλα ένζυμα. Επίσης οι γενετικές θέσεις 196 και 209 είναι πιθανό να αποτελούν τμήμα Alu sequence (εξωγενούς προέλευσης αλληλουχία) και είναι επίσης πολύ σταθερή.

Αλληλουχία που μεταλλάσσεται	Κωδικίο που μεταλλάσσεται	Τύπος μετάλλαξης	Το μεταλασσόμενο κωδικίο περιέχει CpG	Αλληλουχία πλούσια σε C και G (τετράκλωνη έλικα)	Συμμετρία καθρέφτη (τρίκλωνη έλικα)	Επαναλήψεις της κυτοσίνης (C)	Εναλλαγές πουρίνης και πυριμιδίνης (Z-DNA )	Αλληλουχία πλούσια σε A και T (αποελίκωση)	Το μεταλασσόμενο κωδικίο περιέχει TG και CA (εναλλακτική διαμόρφωση)	Η αλληλουχία περιέχει επαναλήψεις GA (κολλώδες DNA)	Η αλληλουχία περιέχει επαναλήψεις δινουκλεοτιδίων (διαμόρφωση φουρκέτας)
------------------------------	---------------------------	------------------	---------------------------------------	--	-------------------------------------	-------------------------------	---	--	--	---	--

### (A) Hotspots σωματικών και γενετικών μεταλλάξεων

150 151 152 153 154 ACA CCC CCG CCC GGC	152 CCG	C>T INS	NAI	NAI	152 CCG CCC	NAI					3*GT
156 157 158 159 160 CGC GTC CGC GCC ATG	158 CGC	G>A C>G  C>T G>T (μόνο σωματικές)	NAI	NAI			NAI				
173 174 175 176 177 GTG AGG CGC TGC CCC	175 CGC	C>T G>A	NAI	NAI			NAI				
211 212 213 214 215 ACT TTT CGA CAT AGT	213 CGA	C>T (κυρίως)  G>A G>T	NAI					NAI			

Αλληλουχία που μεταλλάσσεται	Κωδικίο που μεταλλάσσεται	Τύπος μετάλλαξης	Το μεταλλαγμένο κωδικίο περιέχει CpG	Αλληλουχία πλούσια σε C και G (τετράκλωνη έλικα)	Συμμετρία καθρέφτη (τρίκλωνη έλικα)	Επαναλήψεις της κυτοσίνης (C)	Εναλλαγές πουρίνης και πυριμιδίνης (Z-DNA )	Αλληλουχία πλούσια σε A και T (αποελίκωση)	Το μεταλλαγμένο κωδικίο περιέχει TG και CA (εναλλακτική διαμόρφωση)	Η αλληλουχία περιέχει επαναλήψεις GA (κολλώδες DNA)	Η αλληλουχία περιέχει επαναλήψεις δινουκλεοτιδίων (διαμόρφωση φουρκέτας)
218 219 220 221 222 GTG CCC TAT GAG CCG	220 TAT	A>G (κυρίως)  A>C		NAI		NAI			NAI		
243 244 245 246 247 ATG GGC GGC ATG AAC	245 GGC	G>A (κυρίως)  G>T									
246 247 248 249 250 ATG AAC CGG AGG CCC	248 CGG	G>A C>T	NAI	NAI	248 CGG AGG						
271 272 273 274 275 GAG GTG CGT GTT TGT	273 CGT	G>A C>T	NAI				NAI		NAI		4*GT
279 280 281 282 283 GGG AGA GAC CGGCGC  284 285 286 ACA GAG GAA	282 CGG	C>T	NAI	NAI	282 CGG CGC		NAI				

Αλληλουχία που μεταλλάσσεται	Κωδικίο που μεταλλάσσεται	Τύπος μετάλλαξης	Το μεταλλαγμένο κωδικίο περιέχει CpG	Αλληλουχία πλούσια σε C και G (τετρακλώνη έλικα)	Συμμετρία καθρέφτη (τρίκλωνη έλικα)	Επαναλήψεις της κυτοσίνης (C)	Εναλλαγές πουρίνης και πυριμιδίνης (Z-DNA )	Αλληλουχία πλούσια σε A και T (αποελίκωση)	Το μεταλλαγμένο κωδικίο περιέχει TG και CA (εναλλακτική διαμόρφωση)	Η αλληλουχία περιέχει επαναλήψεις GA (κολλώδες DNA)	Η αλληλουχία περιέχει επαναλήψεις δινουκλεοτιδίων (διαμόρφωση φουρκέτας)
<b>(B) Hotspots μόνο των γεννητικών κυττάρων</b>											
131 132 133 134 135 AAC AAG ATG TTT TGC	133 ATG	T>G T>C						NAI	NAI		
194 195 196 197 198 CTT ATC CGA GTG GAA	196 CGA	C>T	NAI								
207 208 209 210 211 GAT GAC AGA AAC ACT	209 AGA	A>T DEL							NAI	NAI	NAI
233 234 235 236 237 CAC TAC AAC TAC ATG	235 AAC	A>G			235 TAC AAC				NAI		NAI
304 305 306 307 308 ACT AAG CGA GCA CTG	306 CGA	C>T	NAI								
335 336 337 338 339 CGT GAG CGC TTC GAG	337 CGC	G>A (κυρίως) C>T	NAI	NAI							

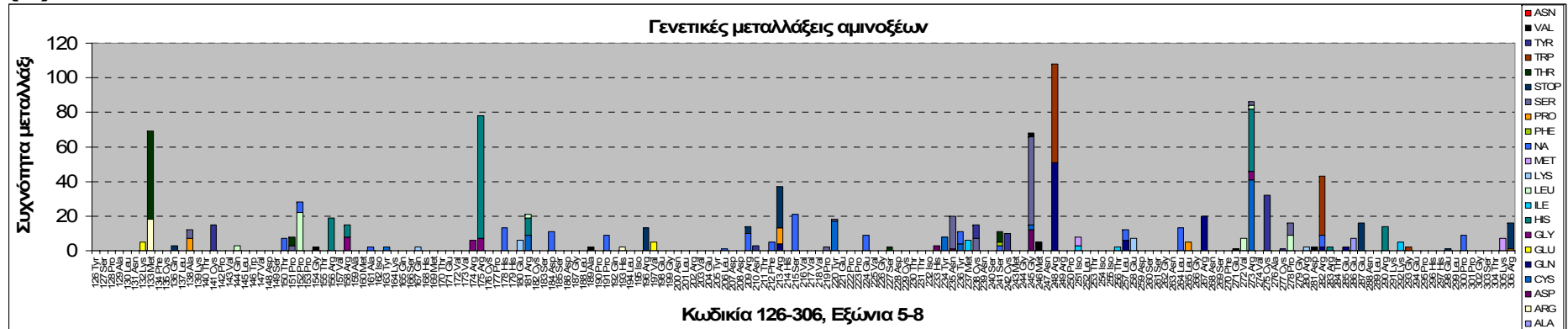
Αλληλουχία που μεταλλάσσεται	Κωδικίο που μεταλλάσσεται	Τύπος μετάλλαξης	Το μεταλλαγμένο κωδικίο περιέχει CpG	Αλληλουχία πλούσια σε C και G (τετράκλωνη έλικα)	Συμμετρία καθρέφτη (τρίκλωνη έλικα)	Επαναλήψεις της κυτοσίνης (C)	Εναλλαγές πουρίνης και πυριμιδίνης (Z-DNA )	Αλληλουχία πλούσια σε A και T (αποελίκωση)	Το μεταλλαγμένο κωδικίο περιέχει TG και CA (εναλλακτική διαμόρφωση)	Η αλληλουχία περιέχει επαναλήψεις GA (κολλώδες DNA)	Η αλληλουχία περιέχει επαναλήψεις δινουκλεοτιδίων (διαμόρφωση φουρκέτας)
<b>(Γ) Hotspots μόνο των σωματικών κυττάρων</b>											
247 248 249 250 251 AAC CGG AGG CCC ATC	249 AGG	G>T		NAI	247 248 C CGG 249 250 AGG CC						
177 178 179 180 181 CCC CAC CAT GAG CGC	179 CAT	C>T A>G				NAI			NAI		NAI
174 175 176 177 178 AGG CGC TGC CCC CAC	176 TGC	G>T (κυρίως) G>A		NAI		NAI			NAI		
155 156 157 158 159 ACC CGC GTC CGC GCC		G>T (κυρίως)  C>T G>A	NAI	NAI			NAI				4*CG

**Αναλύοντας την αλληλουχία του DNA με βάση τις γνώσεις από την εισαγωγή της εργασίας προσπαθούμε να εξηγήσουμε γιατί αυτά τα σημεία μεταλλάσσονται πολύ πιο εύκολα από τα υπόλοιπα**

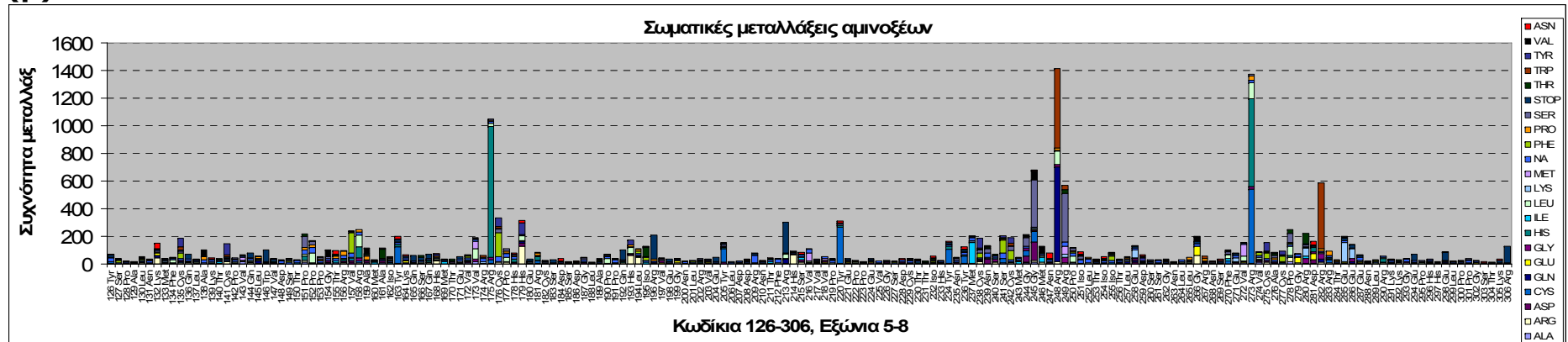
1. Παρατηρούμε ότι η συμμετρία καθρέφτη μπορεί να προκαλέσει το σχηματισμό τρίκλωνης έλικας (triplex).
2. Οι επαναλήψεις δινουκλεοτιδίων GA μπορούν επίσης να οδηγήσουν στο σχηματισμό τρίκλωνης έλικας (triplex) με διαμόρφωση (H – DNA).
3. Οι εναλλαγές πουρίνης και πυριμιδίνης κάνουν την αλληλουχία ευάλωτη στο σχηματισμό αριστερόστροφης έλικας DNA (Z – DNA).
4. Η αλυσίδα που είναι πλούσια σε G:C περιεχόμενο, εν δυνάμει παίρνει εναλλακτική διαμόρφωση τετράκλωνης έλικας (tetraplex).
5. Παρατηρούμε επαναλήψεις των δινουκλεοτιδίων GA και αυτό μπορεί να οδηγήσει στο σχηματισμό κολλώδους DNA (sticky DNA).
6. Το δινουκλεοτίδιο TG, σύμφωνα με την εισαγωγή, παρουσιάζει δομική αμφισταθερότητα και μπορεί να πάρει εναλλακτικές διαμορφώσεις. Το ίδιο και το δινουκλεοτίδιο CA.
7. Επίσης, αλληλουχία πλούσια σε A και T αυτό οδηγεί σε αποελικωση του δίκλωνου μορίου του DNA.
8. Οι επαναλήψεις δι- και τρινουκλεοτιδίων μπορούν να κάνουν ευάλωτη την αλληλουχία στο σχηματισμό φουρκέτας.



(α)



(β)



**Σχήμα 21. Μεταλλαγμένα αμινοξέα σε γενετικό και σποραδικό καρκίνο.** Τα γραφήματα δείχνουν τις μεταλλάξεις των αμινοξέων της πρωτεΐνης του *p53*. Τα χρώματα των αμινοξέων στα δύο σχήματα συμπίπτουν.

Εντοπίζουμε τα αμινοξέα που αντιστοιχούν στα hotspots:

**Σωματικά:** 175 Αργινίνη, 176 Κυστεΐνη, 179 Ιστιδίνη, 213 Αργινίνη, 220 Τυροσίνη, 245 Γλυκίνη, 248 Αργινίνη, 249 Αργινίνη, 273 Αργινίνη και 282 Αργινίνη.

**Γενετικά:** 133 Μεθειονίνη, 152 Προλίνη, 158 Αργινίνη, 175 Αργινίνη, 196 Αργινίνη, 209 Αργινίνη, 213 Αργινίνη, 220 Τυροσίνη, 235 Ασπαραγίνη, 245 Γλυκίνη, 248 Αργινίνη, 273 Αργινίνη, 282 Αργινίνη, 306 Αργινίνη, 337 Αργινίνη.

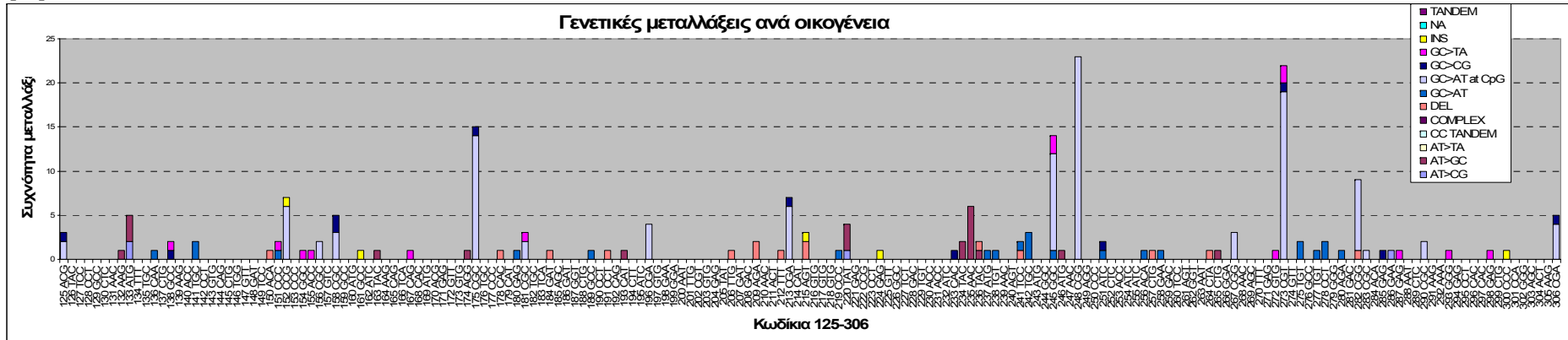
**iMARS για γενετικά:** 152 Προλίνη, 213 Αργινίνη, 235 Ασπαραγίνη, 248 Αργινίνη, 273 Αργινίνη, 282 Αργινίνη.

Παρατηρούμε ότι η μετάλλαξη της Αργινίνης μπορεί να οδηγήσει συχνότερα σε καρκινογένεση. Αυτό το αποτέλεσμα είναι συνισταμένη της μεταλλαξιμότητας της Αργινίνης και της τοξικής επίδρασης των μεταλλάξεών της. Η Αργινίνη, ως παράγοντας προσκόλλησης στο DNA, φαίνεται ότι στα συγκεκριμένα hotspots έχει σπουδαίο ρόλο στη λειτουργικότητα της πρωτεΐνης TP53 και είναι αναντικατάστατη.

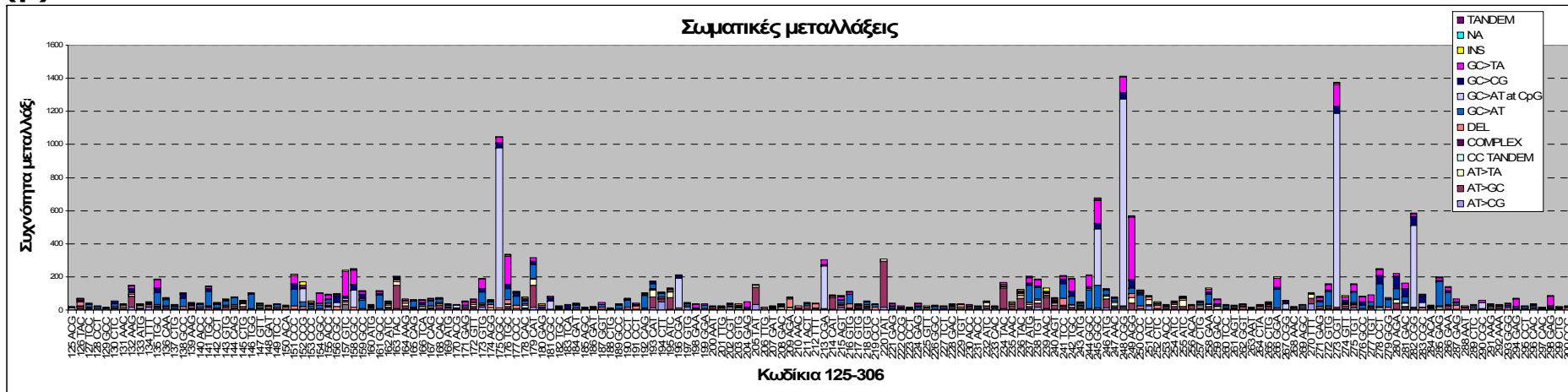
Από τη μελέτη μόνο των κωδικίων της αργινίνης του παρακάτω πίνακα, γίνεται αντιληπτό ότι τα κωδικία που ξεκινούν από CG μεταλλάσσονται στατιστικώς πολύ πιο σημαντικά από αυτά που ξεκινούν με AG και στις γενετικές και στις σωματικές μεταλλάξεις.

<b>Αργινίνη</b>	<b>Γενετικές μεταλλάξεις (282)</b>	<b>Σωματικές μεταλλάξεις (21587)</b>
<b>AGA</b>	5	306
<b>AGG</b>	1	634
<b>CGA</b>	17	703
<b>CGC</b>	52	1673
<b>CGG</b>	35	2059
<b>CGT</b>	22	1456
<b>Σύνολο μεταλλάξεων αργινίνης</b>	132	6840
<b>Δοκιμασία Χ2 μεταξύ κωδικίων που ξεκινούν από AG και CG</b>	0,0000	0,0000

(α)



(β)



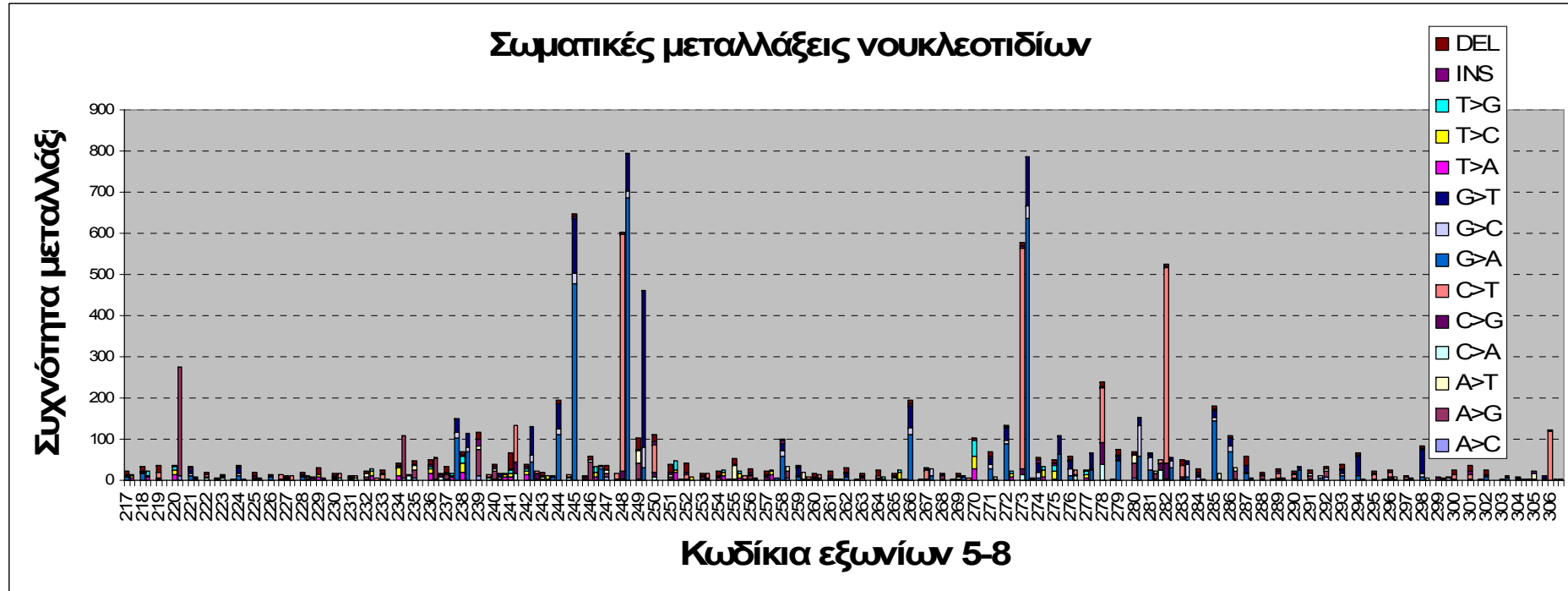
**Σχήμα 22. Τύποι μεταλλάξεων των κωδικίων στην αλληλουχία του p53 στα εξώνια 5-8.** Στα δύο γραφήματα παρατηρούνται στην πλειοψηφία οι ίδιοι τύποι μεταλλάξεων στα αντίστοιχα κωδικία.







(β)



**Σχήμα 23. Απεικόνιση χωριστά των μεταλλάξεων κάθε νουκλεοτιδίου στην αλληλουχία του γονιδίου στα εξώνια 5-8.**

Από τα σχήματα γίνεται φανερό ότι στη συντριπτική πλειοψηφία των περιπτώσεων κάθε ένα νουκλεοτίδιο μεταλλάσσεται με τον ίδιο τρόπο στις γενετικές και σωματικές μεταλλάξεις. Πιθανόν αυτό είναι το τελικό αποτέλεσμα που διαμορφώνεται με την επίδραση τριών παραγόντων:

- ✓ Της αλληλουχίας που βρίσκεται πριν και μετά από το νουκλεοτίδιο και τα πρότυπα που σχηματίζει. Τα πρότυπα αυτά μπορεί να καθιστούν τη δομή ευάλωτη σε επιγενετικές τροποποιήσεις και μεταλλάξεις.
- ✓ Των επιδιορθωτικών μηχανισμών που μπορεί να επιδράσουν σε συγκεκριμένη σημειακή μετάλλαξη.
- ✓ Της τοξικότητας που έχει κάθε μετάλλαξη ξεχωριστά. Δηλαδή πόσο προάγεται η διαδικασία καρκινογένεσης από μία συγκεκριμένη μετάλλαξη.

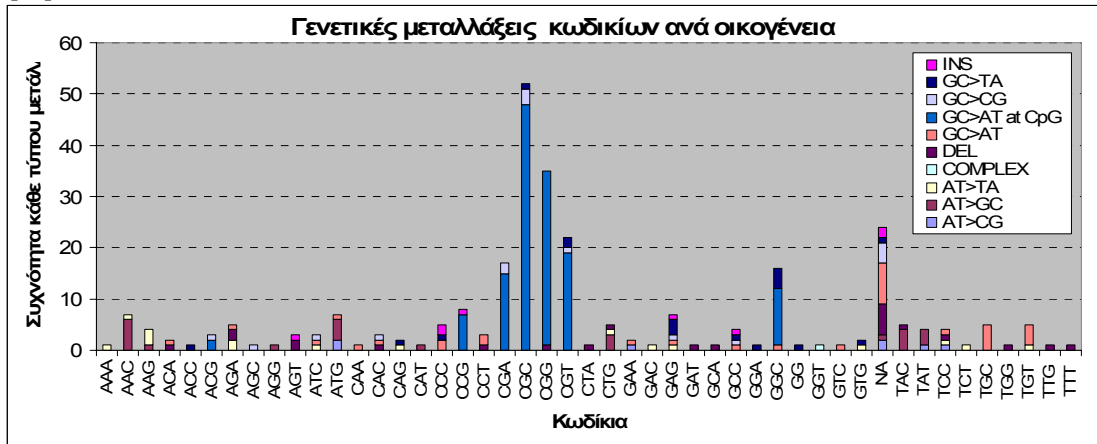
Εφόσον τα νουκλεοτίδια έχουν, κατά πλειοψηφία, ίδιου τύπου μεταλλάξεις στα σωματικά και γενετικά κύτταρα, έχουν ίδιους τους παραπάνω μηχανισμούς μεταλλαξιγένεσης.

Στις περιπτώσεις όπου υπάρχουν αποκλίσεις και οι δύο κυτταρικές σειρές μεταλλάσσονται με διαφορετικό τρόπο, τότε είναι πιθανό να δρουν διαφορετικοί μηχανισμοί μεταλλαξιγένεσης.

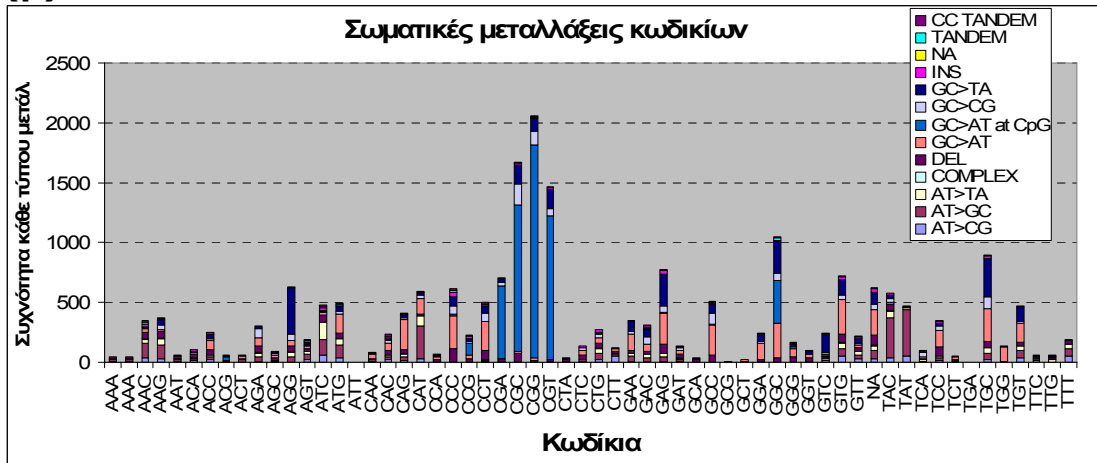
Άρα στα hotspots, η αλληλουχία που τα περιβάλλει τα καθιστά ευάλωτα σε μεταλλάξεις, οι επιδιορθωτικοί μηχανισμοί δεν είναι αποτελεσματικοί και επίσης προάγουν σε μεγάλο βαθμό τη διαδικασία καρκινογένεσης.



(α)



(β)



**Σχήμα 24. Σύγκριση του τύπου των μεταλλάξεων συνολικά για κάθε κωδικίο.**

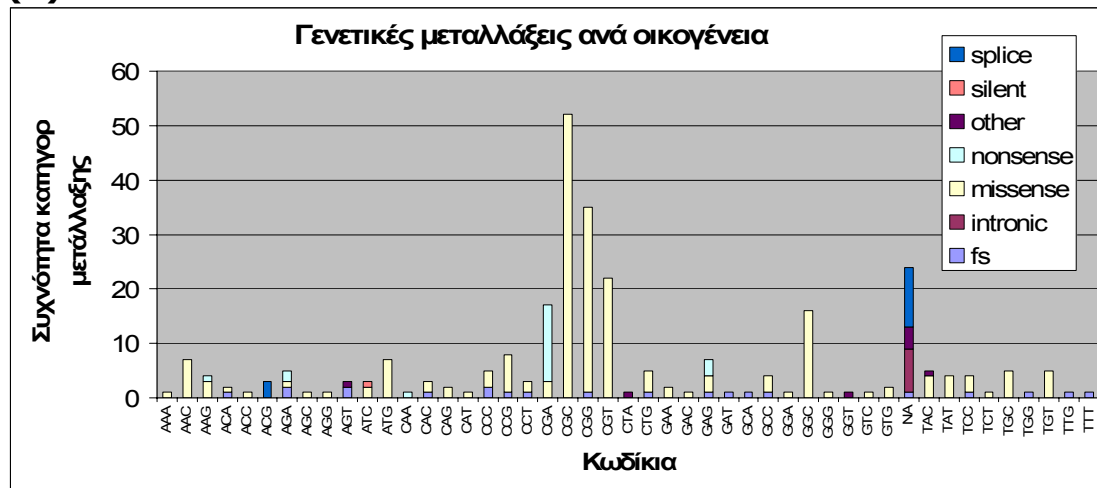
Στα γραφήματα φαίνονται τα κωδικία, ο αριθμός των καταγεγραμμένων μεταλλάξεών τους και η ακριβής μεταβολή των βάσεων τους.

Τα κωδικία που μεταλλάσσονται περισσότερο είναι οι **CGG, CGC, CGT, GGC, CCC, AGG, CGA, GAG, GGC, GTG, TGC** και έχουν κοινό στοιχείο την υψηλή περιεκτικότητα σε CG. Χωρίς όμως αυτό να αποτελεί την απαραίτητη προϋπόθεση καθώς για παράδειγμα τα κωδικία **GCG** και **GCT** παρουσιάζουν μηδενική μεταλλαξιμότητα.

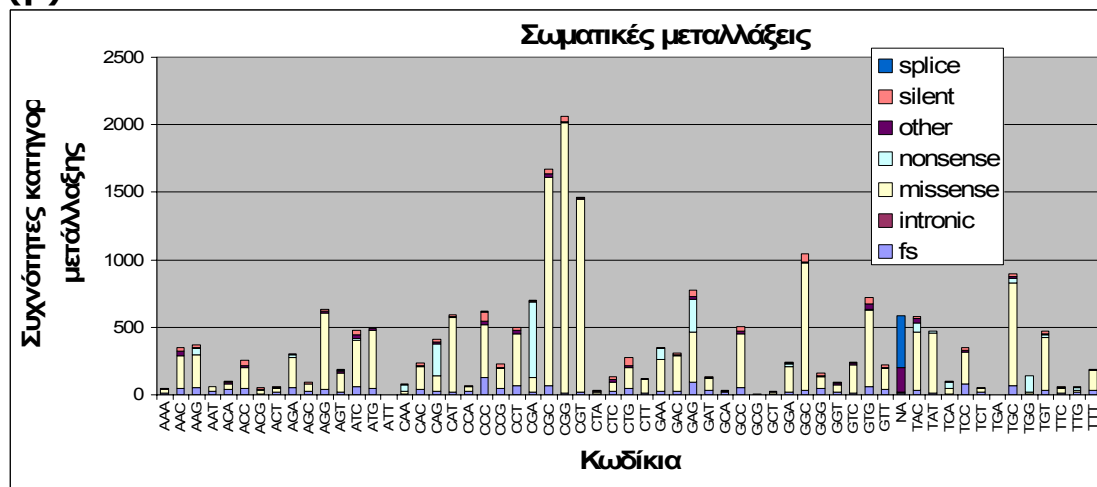
Στα σωματικά κύτταρα, μεθυλίωση και μετατροπή C>T μπορεί να συμβεί όπως παρουσιάζεται σε όλα τα κωδικία που περιέχουν κυτοσίνη ή γουανίνη (πιθανόν μεθυλιώνεται η κυτοσίνη της συμπληρωματικής έλικας), ιδιαίτερα στις σωματικές μεταλλάξεις που περιέχουν μεγάλο αριθμό καταγραφών. Σε όλες, εκτός από την τριπλέτα **TAC**, η οποία κωδικοποιεί το αμινοξύ της τυροσίνης. Ενώ καταγράφονται αρκετά είδη μεταλλάξεων, δεν καταγράφονται μεταλλάξεις μεθυλίωσης της **TAC**, διότι και το κωδικίο **TAT** που

προκύπτει από μεταβολή C>T κωδικοποιεί πάλι την κυτοσίνη και η μετάλλαξη είναι σιωπηρή. Άρα με αυτή την εξαίρεση παρατηρούμε μεγαλύτερο αριθμό G:C>A:T μεταλλάξεων στα σωματικά κύτταρα σε σχέση με τα γενετικά. Πιθανότατα, αυτό είναι χαρακτηριστικό επιγενετικής τροποποίησης της κυτοσίνης (μεθυλίωσης) και προαγωγής των μεταλλάξεών της.

(a)



(β)



**Σχήμα 25. Κατηγορίες των μεταλλάξεων συνολικά ανά κωδικίο.**

Παρατηρείται ότι σε καθορισμένες γενετικές θέσεις στις σωματικές μεταλλάξεις προστίθενται σε ορισμένο ποσοστό σιωπηρές μεταλλάξεις (silent) και άλλου τύπου μεταλλάξεις (other).

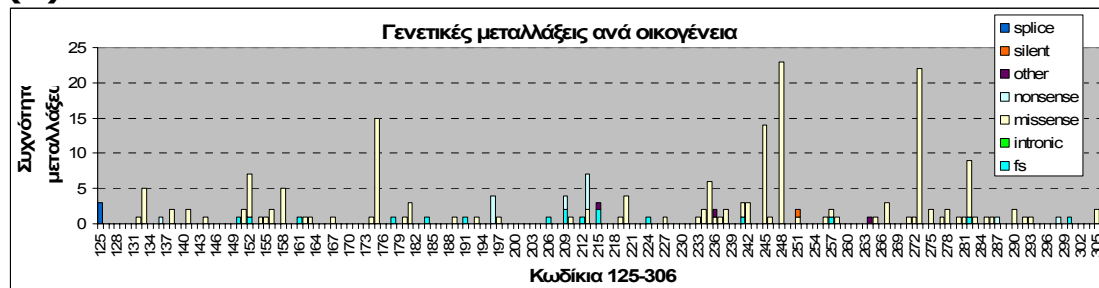
Κάποια κωδικία εξαλλάσσονται περισσότερο σε σωματικό καρκίνο: **AGA, AGG, AGT, ATC, ATG, CAT, CAG, CAC, CCC, CGA, CTG, GAG, GCC, GGC, GTG, GTT, TAC, TAT, TTT, TGT, TGC.** Είναι πιθανό αυτά τα κωδικία:

- ✓ Να είναι επιρρεπή σε επιγενετικές μεθυλιώσεις των σωματικών κυττάρων και κατά συνέπεια ευάλωτα σε μεταλλάξεις
- ✓ Να μην επιδιορθώνονται επαρκώς από τα σωματικά κύτταρα σε σχέση με τα γενετικά
- ✓ Να μεταλλάσσονται ευκολότερα στα σωματικά κύτταρα λόγω της μορφής του DNA σε ευχρωματίνη.

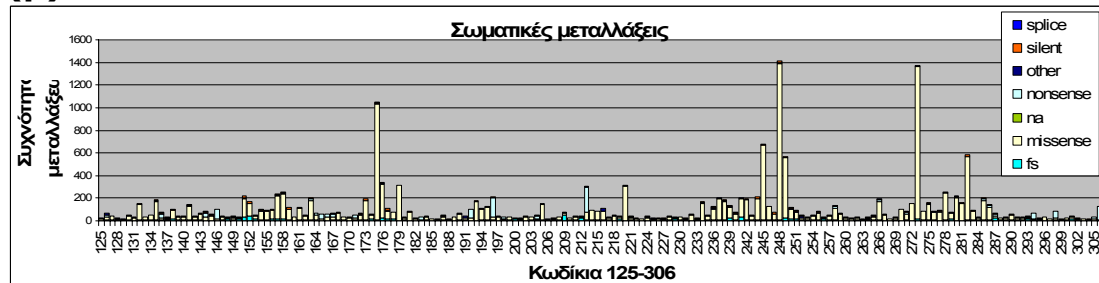
Τα κωδικία που μεταλλάσσονται περισσότερο σε γενετικό καρκίνο είναι: **CTA, GGC**.

Οι σιωπηρές και οι άλλες μεταλλάξεις που απεικονίζονται στο σχήμα 15 (β) είναι ενδεικτικό πιθανότατα γενικευμένης γενετικής αστάθειας.

(α)

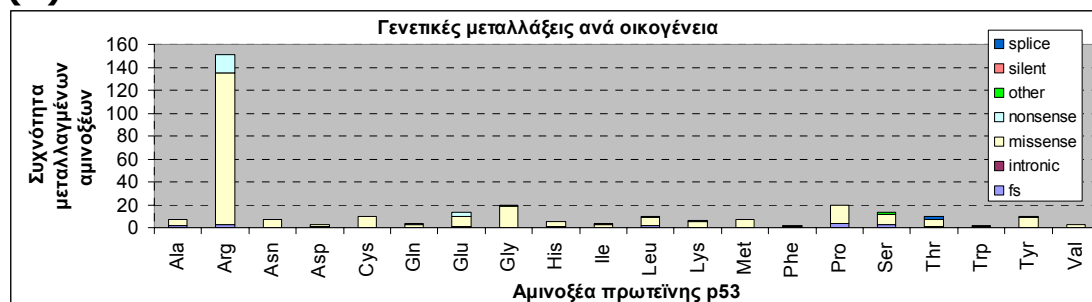


(β)

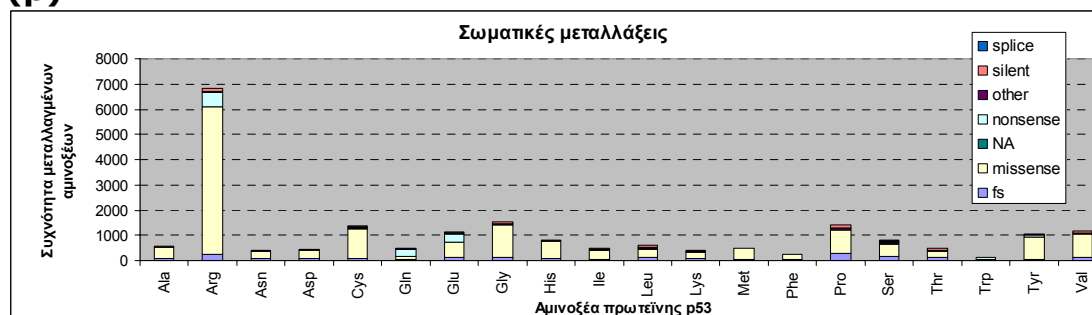


**Σχήμα 26. Κατηγορίες μετάλλαξης των κωδικίων στην αλληλουχία του γονιδίου.** Στο σχήμα αυτό φαίνονται στην αλληλουχία του γονιδίου η κατανομή των μεταλλάξεων ανάλογα με την κατηγορία μετάλλαξης.

(α)



(β)



**Σχήμα 27. Συχνότητα και κατηγορία μετάλλαξης του κάθε αμινοξέως.**

- ✓ Παρατηρούμε ότι τα δύο γραφήματα έχουν πολλές ομοιότητες, που σημαίνει ότι αρκετές μεταλλάξεις σε αντίστοιχα αμινοξέα προκύπτουν με ανάλογη συχνότητα στις δύο σειρές.
- ✓ Το αμινοξύ που μεταλλάσσεται συντριπτικά είναι η αργινίνη και η πλειονότητα των αλλαγών οδηγεί σε missense.
- ✓ Οι σιωπηρές μεταλλάξεις συμβαίνουν στα γεννητικά κύτταρα στην ισολευκίνη, ενώ στα σωματικά συμβαίνουν στην προλίνη, γλυκίνη, βαλίνη και αργινίνη.
- ✓ Η κυστεΐνη, ιστιδίνη, τυροσίνη, βαλίνη και προλίνη μεταλλάσσονται συχνότερα στα σωματικά κύτταρα.

ΓΕΝΕΤΙΚΟΣ ΚΩΔΙΚΑΣ			
<b>Κυστεΐνη</b>	TGC	<b>Ιστιδίνη</b>	CAC
	TGT		CAT
<b>Τυροσίνη</b>	TAC	<b>Βαλίνη</b>	GTA
	TAT		GTC
			GTG
<b>Προλίνη</b>	CCA		GTT
	CCC		
	CCG		
	CCT		

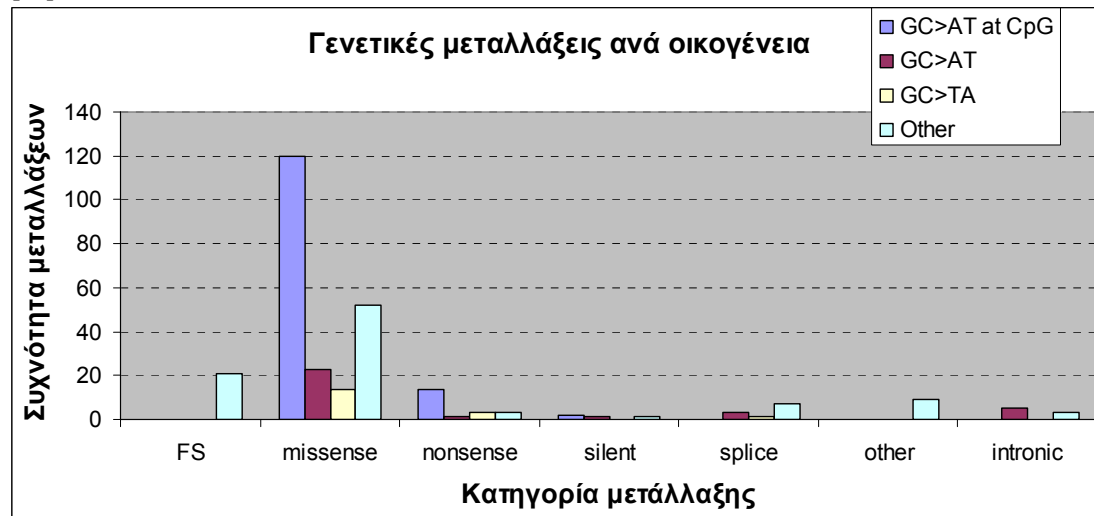
Με τη δοκιμασία  $\chi^2$  έγινε έλεγχος της στατιστικής σημαντικότητας των διαφορών και τα αποτελέσματα παρουσιάζονται στον παρακάτω πίνακα.

	Συχνότητα γενετικών μεταλλάξεων (282)	Συχνότητα σωματικών μεταλλάξεων (21587)	Δοκιμασία $\chi^2$
<b>Αμινοξύ</b>			
<b>Ala</b>	7	575	0,8546
<b>Arg</b>	132	6840	<b>0,0000</b>
<b>Asn</b>	7	417	0,5146
<b>Asp</b>	3	448	0,2425
<b>Cys</b>	10	1371	<b>0,0673</b>
<b>Gln</b>	4	493	0,3416
<b>Glu</b>	13	1127	0,6626
<b>Gly</b>	20	1551	0,9555
<b>His</b>	5	825	<b>0,0820</b>
<b>Ile</b>	4	478	0,3745
<b>Leu</b>	10	625	0,5311
<b>Lys</b>	6	417	0,8161
<b>Met</b>	7	494	0,8328
<b>Phe</b>	2	254	0,4726
<b>Pro</b>	20	1409	0,7215
<b>Ser</b>	13	789	0,4159
<b>Thr</b>	10	476	0,1402
<b>Trp</b>	2	140	0,9004
<b>Tyr</b>	10	1050	0,3264
<b>Val</b>	3	1187	<b>0,0016</b>

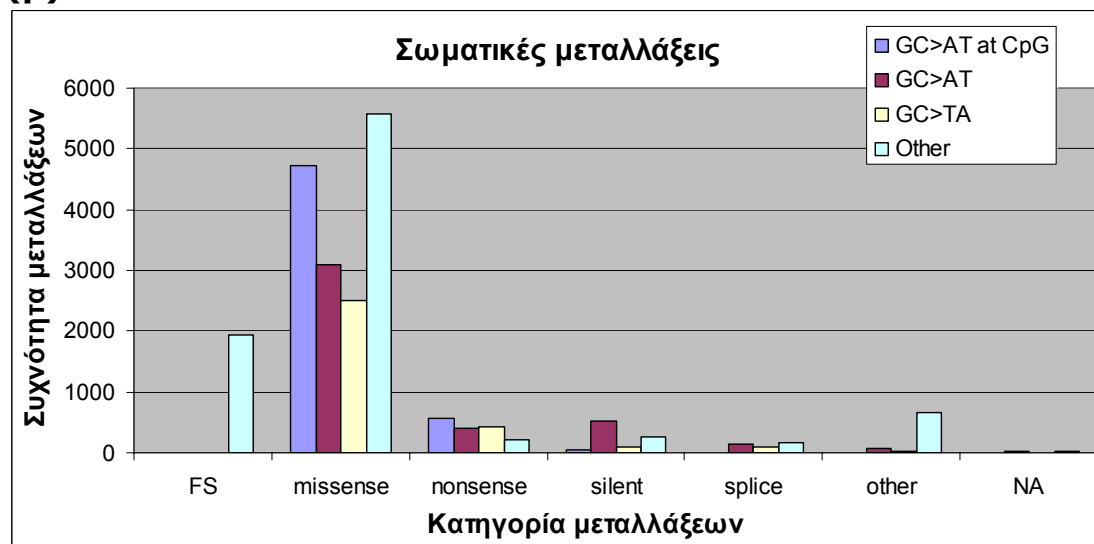
Τα αμινοξέα Βαλίνη και Αργινίνη έχουν διαφορές στατιστικώς πολύ σημαντικές, ενώ τα αμινοξέα Ιστιδίνη και Κυστεΐνη έχουν διαφορές στατιστικώς σημαντικές. Επομένως τα παραπάνω κωδικία είναι επιρρεπή σε επιγενετικές τροποποιήσεις και συνεπακόλουθες μεταλλάξεις που συμβαίνουν στα σωματικά κύτταρα, ιδίως όταν δεν πρόκειται για hotspots.

## Μεταλλάξεις που πιθανόν να σχετίζονται με μεθυλίωση και μεταλλάξεις που δεν σχετίζονται με μεθυλίωση

(α)



(β)



**Σχήμα 28. Κατηγορίες μεταλλαγών σε σχέση με την αστάθεια που έχει αποδοθεί από τη μεθυλίωση της κυτοσίνης και της γουανίνης.**

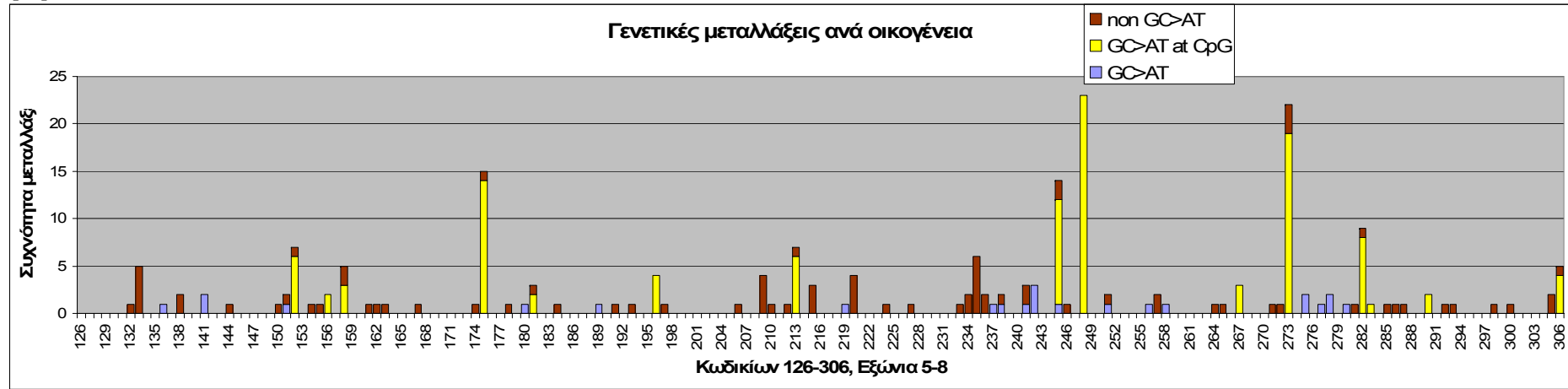
Θεωρούμε ότι οι G:C>A:T μεταλλάξεις προέρχονται από μεθυλίωση της κυτοσίνης. Παρατηρούμε ότι στα σωματικά κύτταρα αυξάνονται οι μεταλλάξεις G:C>A:T που οφείλονται σε non CpG μεθυλίωση. Επίσης αυξάνονται οι G:C>T:A μεταλλάξεις που οφείλονται σε πιθανή μεθυλίωση της γουανίνης ή σε πιθανή μεθυλίωση της κυτοσίνης της συμπληρωματικής αλυσίδας.

Στην κατηγορία των παρανοηματικών (missense) παρατηρούνται σημαντικές διαφορές. Στα σωματικά κύτταρα αυξάνουν οι G:C>A:T,

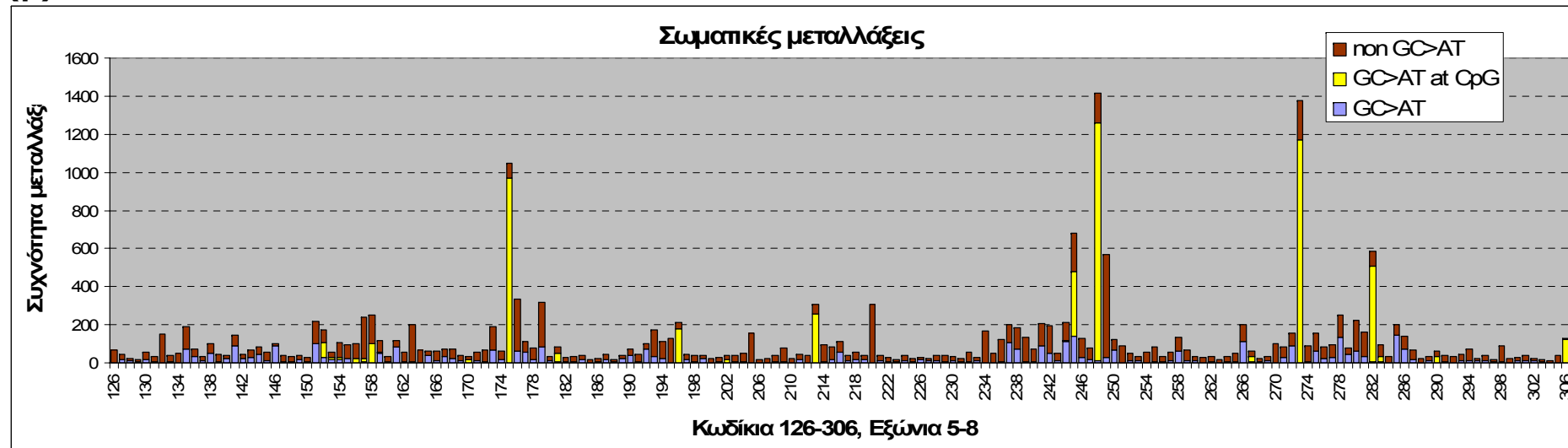
G:C>T:A και Other καταγραφές. Άρα αυξάνονται οι μεταλλάξεις που δεν οφείλονται σε μεθυλίωση του CpG δινουκλεοτιδίου.

Επίσης, είναι αξιοσημείωτο το γεγονός ότι περισσότερες καταγραφές καρκινογένεσης υπάρχουν στις παρανοηματικές μεταλλάξεις (missense) και όχι στις τερματισμού (nonsense) και μετακίνησης του μεταγραφικού πλαισίου (frameshift). Αυτό είναι οξύμωρο διότι οι μεταλλάξεις τερματισμού και μετακίνησης του μεταγραφικού πλαισίου προκαλούν μεγαλύτερες διαταραχές στη λειτουργικότητα της πρωτεΐνης από τις παρανοηματικές. Ίσως τα κύτταρα να έχουν επιδιορθωτικούς μηχανισμούς που δεν επιτρέπουν τη συχνή δημιουργία τέτοιων σημαντικών βλαβών, ενώ οι παρανοηματικές μεταλλάξεις μπορούν να διαφύγουν της επιδιόρθωσης. Ή μπορεί τα κύτταρα να «χάνουν» από τις σιωπηρές μεταλλάξεις σημεία ματίσματος.

(α)



(β)





### **Σχήμα 29. Τύποι μεταλλάξεων που πιθανώς σχετίζονται ή όχι με τη μεθυλίωση της κυτοσίνης.**

Παρατηρούμε ότι δεν είναι όλα τα CpG νουκλεοτίδια hotspots. Υπερτερούν οι μεθυλιωτικές μεταλλάξεις σε CpG δινουκλεοτίδια μόνο στις γενετικές μεταλλάξεις. Στις σωματικές μεταλλάξεις δεν υπερτερούν πάντα οι CpG μεταλλάξεις στα CpG κωδίκια.

Στη γενετική βάση δεδομένων, το σύνολο των CpG μεταλλάξεων είναι 136. Παρατηρούμε ότι είναι αρκετά εντοπισμένες σε γενετικές θέσεις. Στις συγκεκριμένες γενετικές θέσεις είναι πιο εύκολο λόγω δομής του DNA να συμβεί CpG μετάλλαξη. Επομένως, οι CpG μεταλλάξεις προκαλούνται ευκολότερα από τις κυτταρικές διαιρέσεις και το μικροπεριβάλλον των γενετικών κυττάρων.

Το σύνολο των non-CpG μεταλλάξεων είναι 146. Παρατηρούμε ότι είναι περισσότερο διασκορπισμένες. Καταλαμβάνουν περισσότερες γενετικές θέσεις και δε σχηματίζουν έντονα hotspots, σε αντίθεση με τις CpG μεταλλάξεις. Στις non-CpG ξεχωρίζουν οι εξής θέσεις: 133, 209, 220, 235.

Όμως στις σωματικές μεταλλάξεις καταγράφονται περισσότερες θέσεις μετάλλαξης του δινουκλεοτιδίου CpG. Τα CpG δινουκλεοτίδια που μεταλλάσσονται στα γεννητικά κύτταρα συναντώνται και στα σωματικά κύτταρα. Στα σωματικά κύτταρα λόγω του μεγάλου αριθμού των εγγραφών υπάρχουν περισσότερες θέσεις μετάλλαξης του CpG.

Τα κωδίκια 82, 125, 156, 181, 196, 267, 290 και 342 περιέχουν θέσεις CpG. Σ' αυτές υπερισχύουν οι μεταλλάξεις μεθυλίωσης και στα γενετικά και στα σωματικά κύτταρα. Άρα οι κυτοσίνες των CpG είναι μεθυλιωμένες. Παρόλα αυτά δεν αποτελούν hotspots.

Τα δύο γραφήματα έχουν όμοια κατανομή και φαίνεται να υπάρχουν τα ίδια επηρεπή σημεία μετάλλαξης. Παρατηρούμε ότι στα γενετικά κύτταρα το ποσοστό μεταλλάξεων από μεθυλίωση είναι σαφώς μεγαλύτερο σε σχέση με τα σωματικά κύτταρα. Άρα οι μεταλλάξεις οφειλόμενες σε μεθυλίωση είναι πιο εύκολο να μεταφερθούν από γενιά σε γενιά ή να συμβούν στις διαδικασίες γαμετογένεσης και γονιμοποίησης.

Τα πιο έντονα hotspots που προκαλούνται λόγω μεγάλου αριθμού μεταλλάξεων μεθυλίωσης εντοπίζονται στα κωδίκια 175, 213, 245, 248, 273, 282. Τα hotspots που υπάρχουν μόνο στα γενετικά κύτταρα και δεν οφείλονται σε μεταλλάξεις μεθυλίωσης είναι στα κωδίκια 133 και 235. Τα hotspots που υπάρχουν μόνο στα σωματικά κύτταρα και δεν οφείλονται σε μεθυλίωση είναι 157, 176, 179, 220 και 249.

**4. Οι αναλύσεις που γίνονται στη συνέχεια λαμβάνουν υπόψη το σύνολο των γενετικών μεταλλάξεων (1424) και όχι μόνο τις μεταλλάξεις ανά οικογένεια (282). Οι αναλύσεις γίνονται μόνο στη γενετική βάση μεταλλάξεων διότι εκεί μόνο είναι καταγεγραμμένες οι παράμετροι που μελετούνται.**

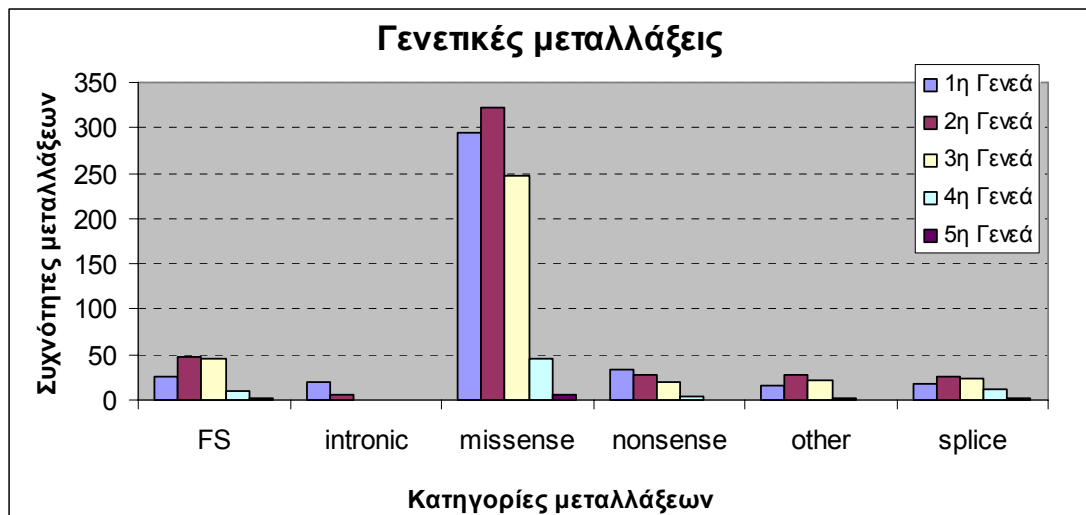
### **Διεισδυτικότητα των γενετικών μεταλλάξεων**

Η ανάλυση των γενεών είναι εφικτή μόνο στα γενετικά κύτταρα, διότι εκεί μόνο είναι καταγεγραμμένη αυτή η παράμετρος. Η πρώτη γενεά είναι τα άτομα μεγαλύτερης ηλικίας και η πέμπτη γενεά τα άτομα μικρότερης ηλικίας.



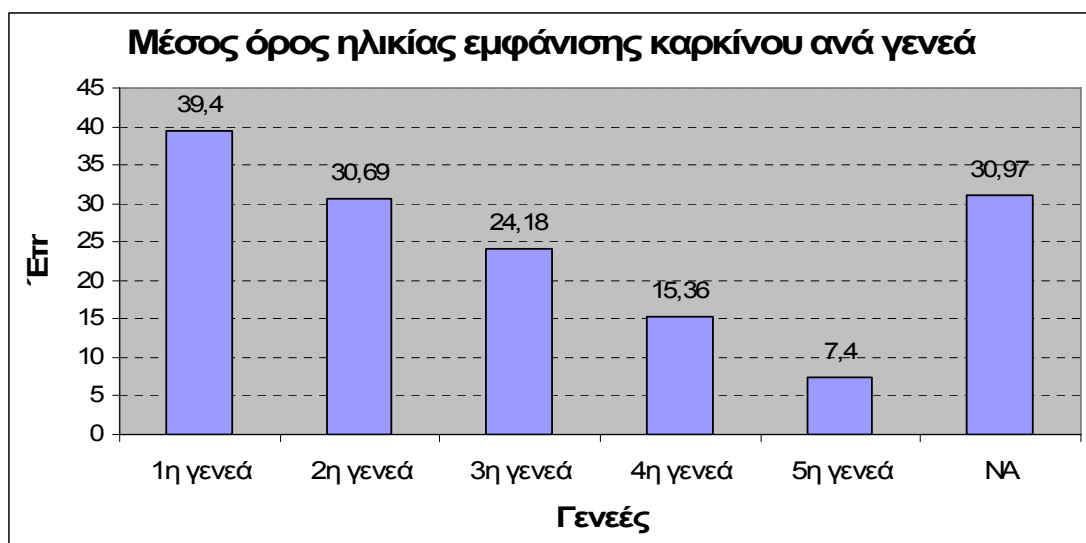
**Σχήμα 30. Συχνότητα μεταλλάξεων ανά γενεά.**

Όπως παρατηρούμε τα νεαρά άτομα της 5<sup>ης</sup> γενεάς ακόμη δεν έχουν εκδηλώσει μεγάλο αριθμό νεοπλασιών και είναι αδύνατη η εξαγωγή συμπερασμάτων. Τα άτομα της 4<sup>ης</sup> γενεάς έχουν αρχίσει να εμφανίζουν νεοπλασίες και θα έχουν την τάση να φτάσουν τα επίπεδα της 3<sup>ης</sup> και 2<sup>ης</sup> γενεάς.



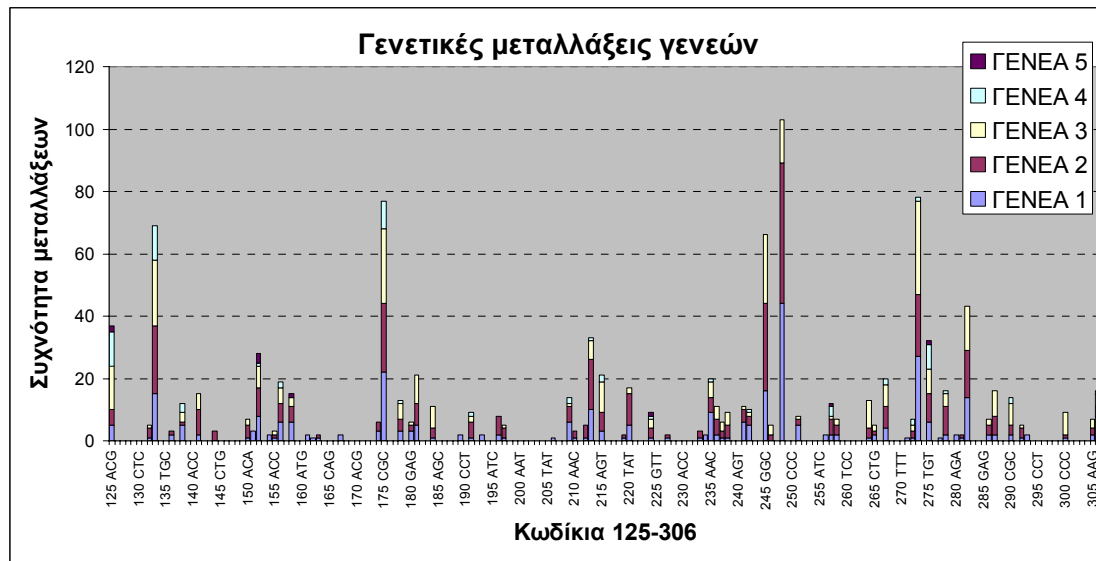
**Σχήμα 31. Κατηγορίες μεταλλάξεων ανά γενεά.**

Πιο καλά μελετημένες είναι η 1<sup>η</sup>, 2<sup>η</sup> και 3<sup>η</sup> γενεά. Η πλειονότητα των μεταλλάξεων είναι παρανοηματικές (missense), ενώ η συχνότητα των υπολοίπων είναι κάτω από 50 μεταλλάξεις. Η πρώτη γενεά φαίνεται έχει ποσοστά χαμηλότερα από της 2<sup>ης</sup> και 3<sup>ης</sup>, εκτός από τις μεταλλάξεις των ενδονίων (intronic) και τερματισμού (nonsense).

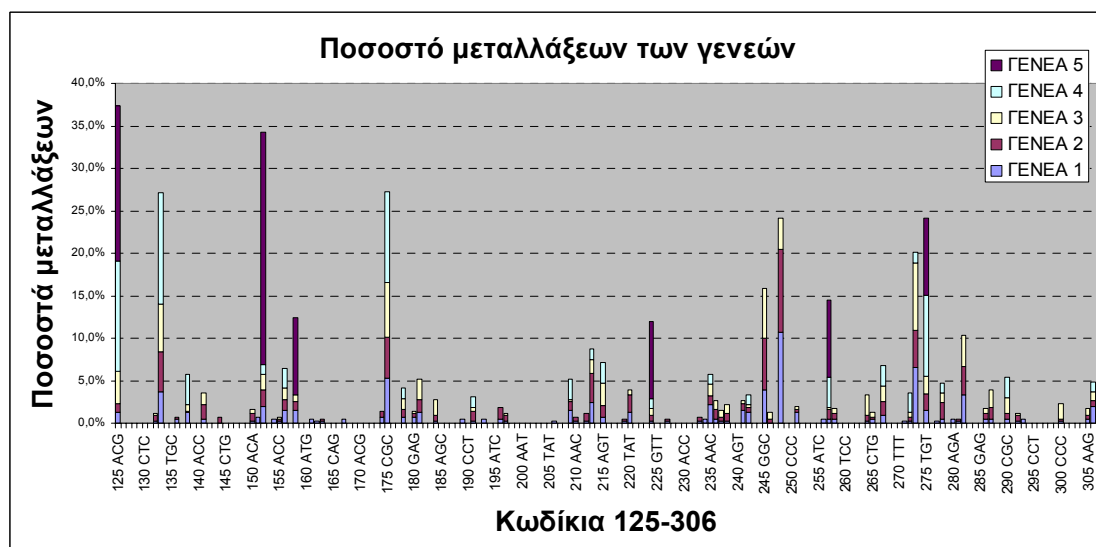


**Σχήμα 32. Μέσος όρος ηλικίας εμφάνισης καρκίνου ανά γενεά.**

Ο συντελεστής συσχέτισης του Pearson ( $r$ ) δίνει ως αποτέλεσμα 0,909 για τη συσχέτιση πλήθους καταγραφών ανά γενεά και μέση ηλικία εμφάνισης καρκίνου. Αυτό σημαίνει ότι υπάρχει πολύ ισχυρή θετική συσχέτιση μεταξύ τους. Άρα όσο μεγαλώνει σε ηλικία ο πληθυσμός με βλάβη του 53, τόσο αυξάνονται τα περιστατικά καρκινογένεσης με γραμμική σχέση, ενώ στο γενικό πληθυσμό η σχέση αυτή απεικονίζεται ως γεωμετρική. [132]



**Σχήμα 33. Συχότητα μετάλλαξης κωδικίων στην αλληλουχία του γονιδίου ανά γενεά.**



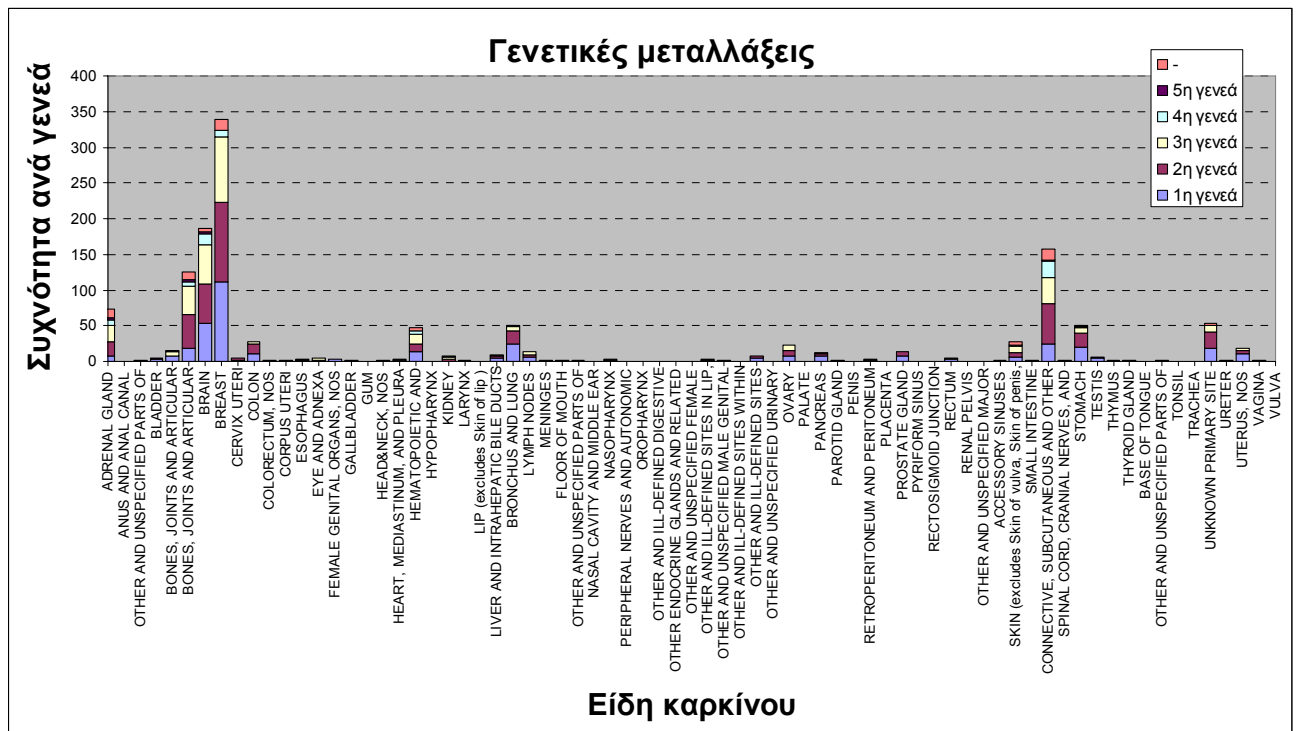
**Σχήμα 34. Συσχέτιση ορισμένων κωδικονίων με ηλικία εμφάνισης νεοπλασματος.**

Τα ποσοστά αποτελούν τον λόγο της συχνότητας μιας μετάλλαξης προς τις συνολικές μεταλλάξεις της γενεάς που ανήκει επί τοις εκατό.

Το σχήμα 24 είναι διευκρινιστικό των hotspots που υπάρχουν. Δεχόμαστε αυθαίρετα ως hotspots κωδικία με συχνότητα μετάλλαξης πάνω από 20.

Τα κωδικία που μεταλλάσσονται πρώτα ή προκαλούν πρώτα καρκινογένεση ανήκουν στην 5<sup>η</sup> γενεά. Ταυτόχρονα παρουσιάζουν μεγάλη διεισδυτικότητα επειδή συναντώνται σε όλες τις γενεές. Αυτά που μεταλλάσσονται δεύτερα ή προκαλούν πιο αργή εμφάνιση νεοπλασμάτων βρίσκονται στην 4<sup>η</sup> γενεά και αυτά που μεταλλάσσονται τρίτα βρίσκονται στις υπόλοιπες γενεές.





**Σχήμα 35. Είδη καρκίνου που εμφανίζονται ανά γενεά.**

Παρατηρούμε ότι δεν υπάρχει σαφής προτίμηση συγκεκριμένης ηλικίας-γενεάς σε συγκεκριμένο είδος καρκίνου. Τα νεοπλάσματα κατανέμονται σε όλα τα όργανα. Νεοπλάσματα με τη μεγαλύτερη συχνότητα είναι των επινεφριδίων, των οστών και αρθρώσεων, του εγκεφάλου, του μαστού, του συνδετικού ιστού και της αγνώστου πρωτοπαθούς εστίας. Αρκετά συχνά είναι επίσης του αιμοποιητικού, του πνεύμονα και του στομάχου.

Νεοπλάσματα που εμφανίζονται σε νεαρότερη ηλικία:

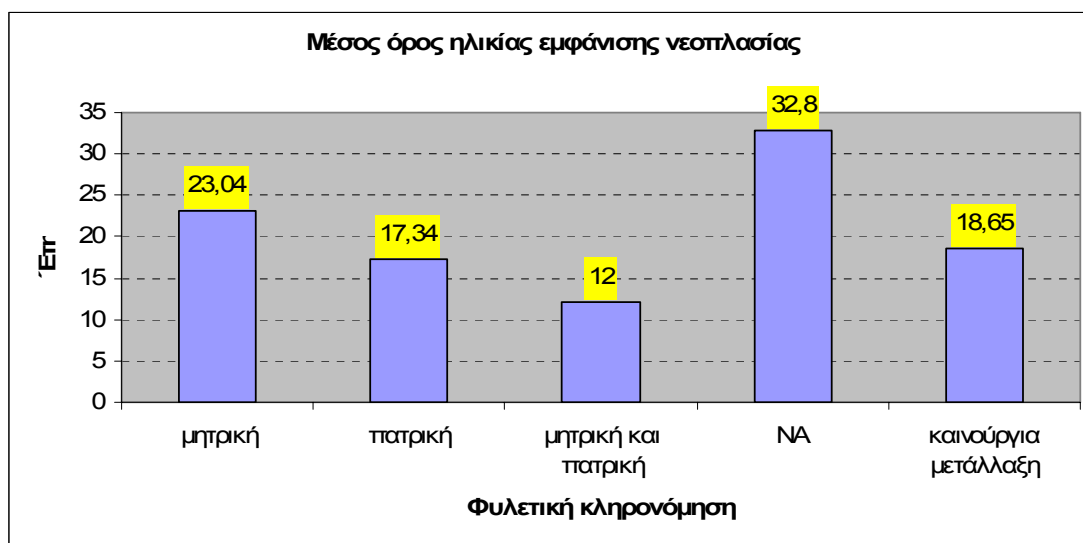
- ✓ Adrenal Gland,
- ✓ Bones,
- ✓ Brain,
- ✓ Breast,
- ✓ Hematopoietic,
- ✓ Connective.

Νεοπλάσματα που εμφανίζονται σε μεγαλύτερη ηλικία:

- ✓ Bronchus and Lung,
- ✓ Skin,
- ✓ Ovary,
- ✓ Uterus,
- ✓ Stomach.

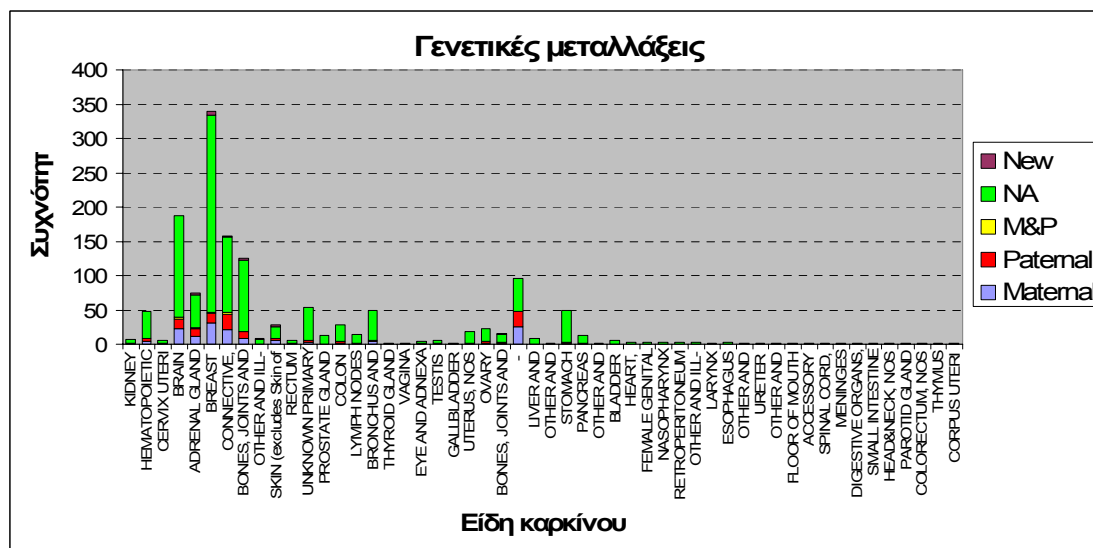
Στα νεοπλάσματα μεγαλύτερης ηλικίας φαίνεται ότι είναι σημαντική η επίδραση περιβαλλοντικών παραγόντων.

## Συσχέτιση φυλετικής κληρονομής και γενετικών μεταλλάξεων



**Σχήμα 36. Μέσος όρος ηλικίας εμφάνισης νεοπλασίας στα γενετικά κύτταρα.**

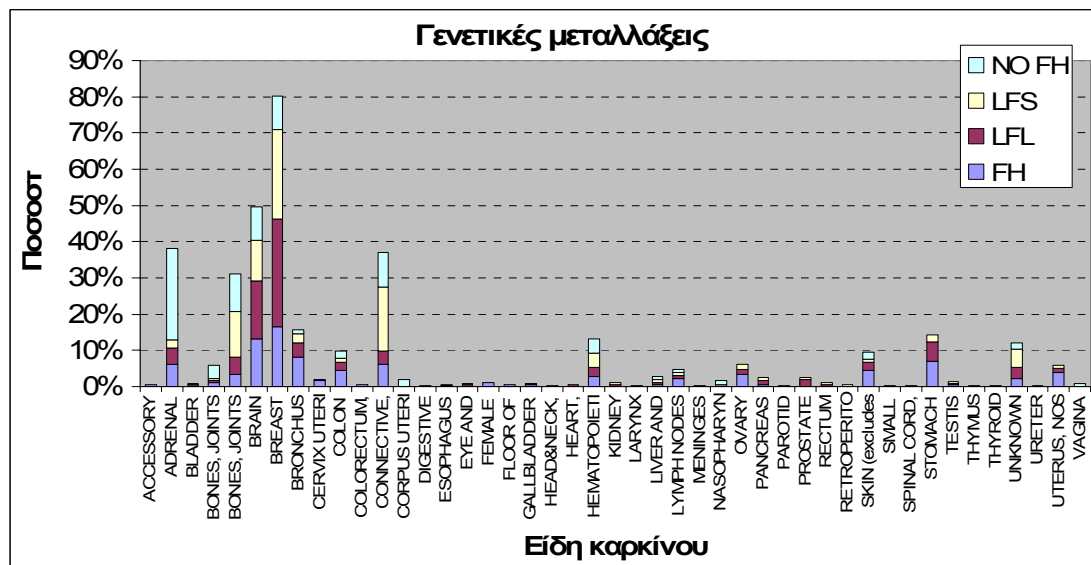
Ο μέσος όρος αυτών που δεν γνωρίζουμε (NA) είναι υψηλότερος και με τις περισσότερες συνολικές εγγραφές. Αυτό αποτελεί έναν έμμεσο δείκτη ότι τα όρια θα τείνουν σε υψηλότερους μέσους όρους. Αν η μετάλλαξη κληρονομείται και από τους δύο γονείς τότε εμφανίζεται νωρίτερα η νεοπλασία. Οι μεταλλάξεις που κληρονομούνται από τον πατέρα έχουν πρωιμότερη εκδήλωση νεοπλασματος.



**Σχήμα 37. Συχνότητα ειδών καρκίνου ανάλογα με την φυλετική κληρονομία.**

Παρατηρούμε ότι δεν υπάρχει τάση εμφάνισης συγκεκριμένου νεοπλασματος ανάλογα με την προέλευση του μεταλλαγμένου *p53* γονιδίου (πατρική ή μητρική).

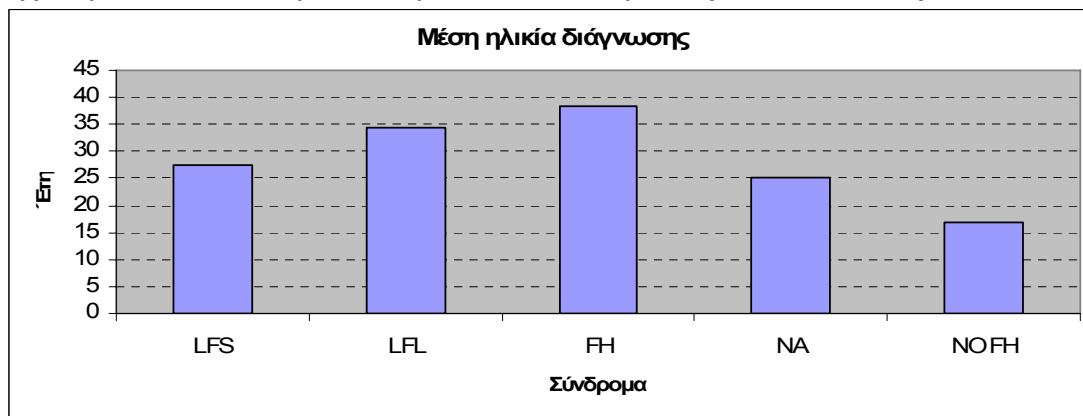
## Είδη γενετικών συνδρόμων και συσχέτιση με καρκινογένεση



**Σχήμα 38. Ποσοστά μεταλλάξεων κάθε γενετικού συνδρόμου για κάθε είδος καρκίνου.**

Τα ποσοστά υπολογίζονται ως πηλίκo των καταγραφών ενός νεοπλασματος δια του συνόλου των καταγραφών όλων των νεοπλασμάτων για το συγκεκριμένο σύνδρομο.

Παρατηρούμε ότι υπάρχει προδιάθεση ανάπτυξης όγκου των επινεφριδίων (Adrenal) στους φορείς μεταλλάξεων χωρίς οικογενειακό ιστορικό (No Family History). Αντίθετα παρατηρούμε ότι η ανάπτυξη καρκίνου του στομάχου και των ωοθηκών σχετίζονται πάντα με οικογενειακό ιστορικό (LFS, LFL, FH).

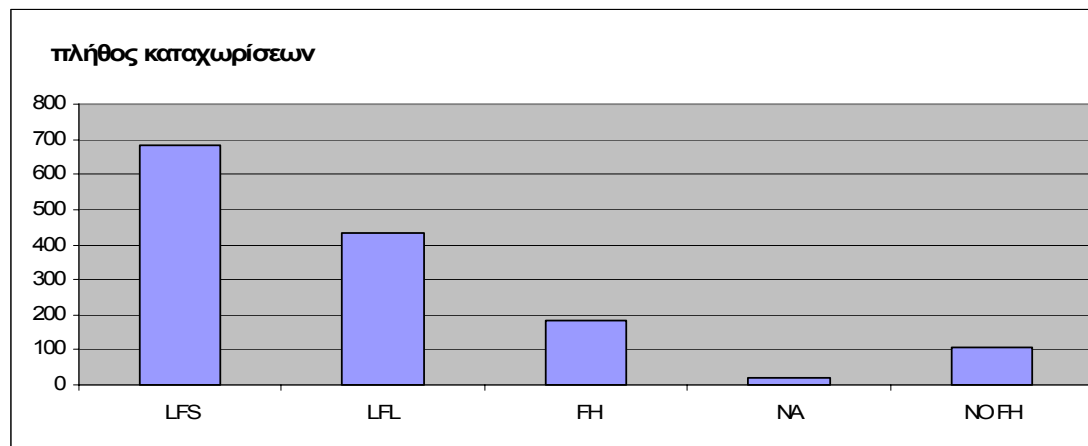


**Σχήμα 39. Ηλικία διάγνωσης του καρκίνου σε σχέση με τα γενετικά σύνδρομα.**

Οι πιο επιθετικές μεταλλάξεις με σειρά ανάλογα με την ηλικία εμφάνισης είναι Χωρίς Ιστορικό > Li Fraumeni Σύνδρομο > Li Fraumeni Like Σύνδρομο > Με οικογενειακό ιστορικό (No FH > LFS > LFL > FH). Οι LFL μεταλλάξεις και οι FH είναι πιο ήπιες από τις LFS.



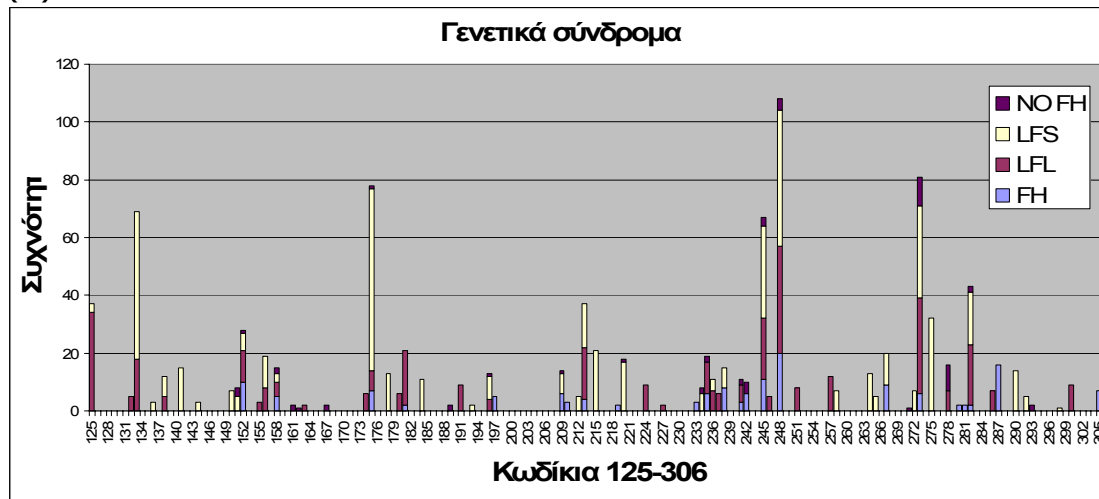
Αντίθετα οι μεταλλάξεις NO FH προκαλούν γρήγορη έναρξη των νεοπλασιών και δεν αφήνουν πολλά περιθώρια για γέννηση απογόνων.



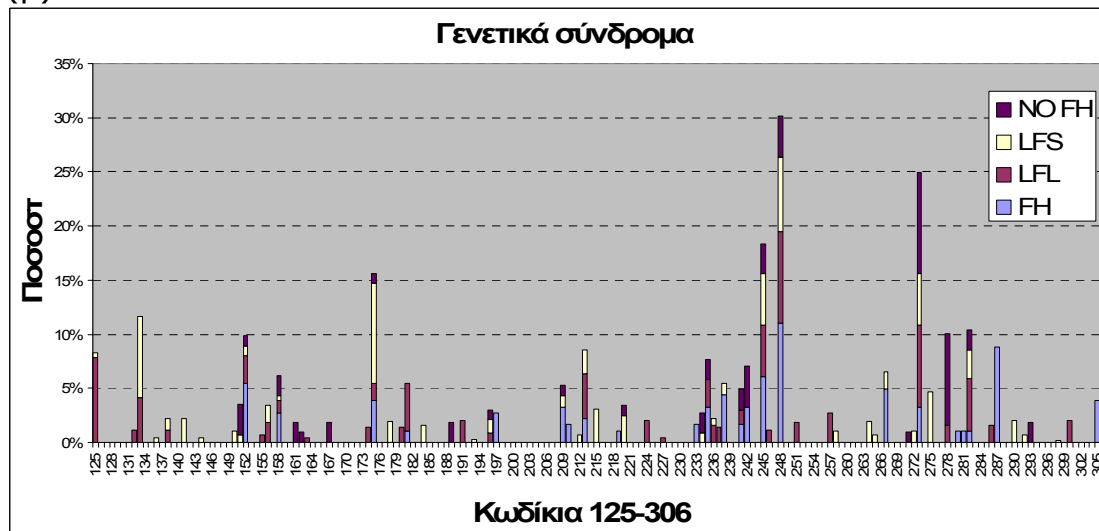
**Σχήμα 40. Πλήθος καταχωρίσεων στη γενετική βάση για το κάθε είδος συνδρόμου.**

Παρατηρούμε ότι μπορούμε να συγκρίνουμε τις κατηγορίες εκτός από την NA που περιέχει πολύ λίγες καταγραφές. Το NA μπορεί να αγνοηθεί διότι έχει μόνο 19 καταγραφές ενώ όλα τα υπόλοιπα πάνω από 100. Το NA σημαίνει απροσδιόριστο από τη βιβλιογραφία.

(α)



(β)



**Σχήμα 41. Μεταλλάξεις κάθε συνδρόμου στην αλληλουχία του γονιδίου.** Στο σχήμα (α) καταγράφονται οι απόλυτοι αριθμοί των μεταλλάξεων. Στο σχήμα (β) το ποσοστό αφορά το πηλίκο του αριθμού των μεταλλάξεων ενός συνδρόμου σε συγκεκριμένο κωδικίο προς τις συνολικές μεταλλάξεις του συνδρόμου.

Οι δύο κυριότερες μεταλλάξεις του NO FH είναι τα κωδικία 273 και 278. Οι υπόλοιπες καταγραφές είναι 151, 158, 161, 167, 189, 241, 242, 245, 248, 273, 282 και 293. Παρατηρούμε σε συνδυασμό με το σχήμα 39 ότι οι μεταλλάξεις σε αυτές τις θέσεις είναι οι πιο επιθετικές και προκαλούν τη γρηγορότερη εμφάνιση νεοπλασμάτων. Σε συνδυασμό με το γράφημα 33 προκύπτει ότι μεταλλάξεις στα παραπάνω κωδικία προδιαθέτουν σε μεγάλο βαθμό στην εμφάνιση όγκων των επινεφριδίων και των οστών.

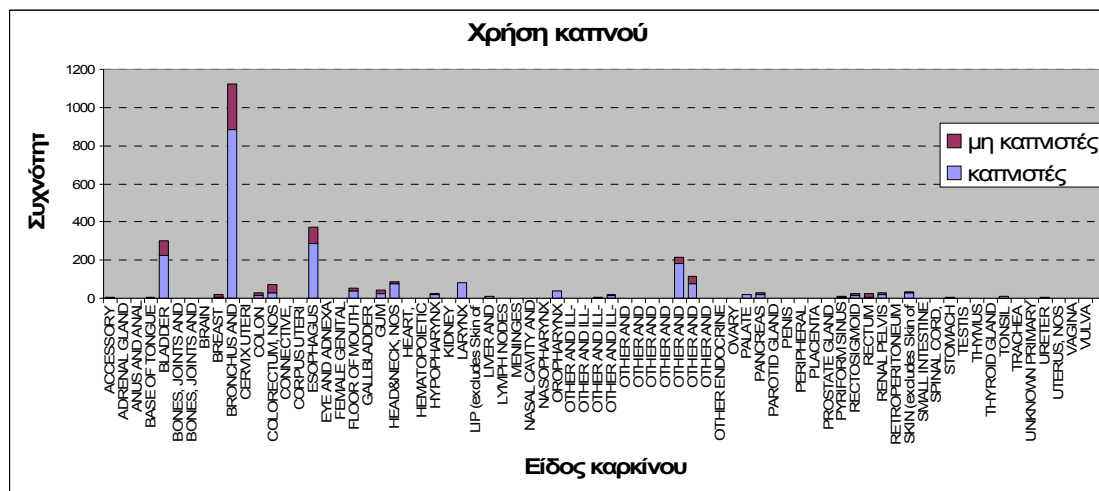
Μεταλλάξεις στα κωδικία 125, 133, 175, 213, 248, προκαλούν κυρίως την εμφάνιση των συνδρόμων LFS και LFL. Σε συνδυασμό με το σχήμα 38 διαπιστώνουμε ότι οι μεταλλάξεις των LFS και LFL σχετίζονται με τη δημιουργία καρκίνου του μαστού, του εγκεφάλου, του πνεύμονα, του συνδετικού ιστού και του στομάχου.

Αναλύεται η αλληλουχία των hotspot μεταλλάξεων χωρίς ιστορικό που δεν έχουν αναλυθεί σε προηγούμενους πίνακες.

Αλληλουχία που μεταλλάσσεται	Κωδικό που μεταλλάσσεται	Τύπος μετάλλαξης	Το μεταλασσόμενο κωδικό περιέχει CpG	Αλληλουχία πλούσια σε C και G (τετράκλωνη έλικα)	Συμμετρία καθρέφτη (τρίκλωνη έλικα)	Επαναλήψεις της κυτοσίνης (C)	Εναλλαγές πουρίνης και πυριμιδίνης (Z-DNA)	Αλληλουχία πλούσια σε A και T (αποελίκωση)	Το μεταλασσόμενο κωδικό περιέχει TG και CA (εναλλακτική διαμόρφωση)	Η αλληλουχία περιέχει επαναλήψεις GA (κολλώδες DNA)	Η αλληλουχία περιέχει επαναλήψεις δινουκλεοτιδίων (διαμόρφωση φουρκέτας)
<b>(ΣΤ) Hotspots των μεταλλάξεων χωρίς ιστορικό</b>											
149 150 151 152 153 TCC ACA CCC CCG CCC 154 GGC	151 CCC	C>T C>A		NAI		NAI					
156 157 158 159 160 CGC GTC CGC GCC ATG	158 CGC	C>G G>A	NAI	NAI			NAI				4*CG
165 166 167 168 169 CAG TCA CAG CAC ATG	167 CAG	C>A							NAI	NAI	5*CA
239 240 241 242 243 AAC AGT TCC TGC ATG 244 GGC	241 TCC	C>T DEL		NAI							
	242 TGC	G>A		NAI							
276 277 278 279 280 GCC TGT CCT GGG AGA	278 CCT	C>T		NAI							

**5. Τα παρακάτω γραφήματα αναφέρονται μόνο στη βάση σωματικών μεταλλάξεων επειδή μόνο σε αυτή είναι καταχωρημένες οι μελετούμενες μεταβλητές. Επειδή υπάρχει μεγάλο ποσοστό των καταγραφών της σωματικής βάσης δεδομένων που δεν γνωρίζουμε τις συνήθειες και τις έξεις τους, τα αποτελέσματά μας προκύπτουν μόνο από ένα μικρό τμήμα της βάσης δεδομένων. Αυτό σημαίνει ότι τα αποτελέσματά μας δεν ισχύουν για ολόκληρη τη βάση δεδομένων.**

### **Συσχέτιση επιδράσεων περιβάλλοντος και εμφάνισης νεοπλασμάτων στα σωματικά κύτταρα**



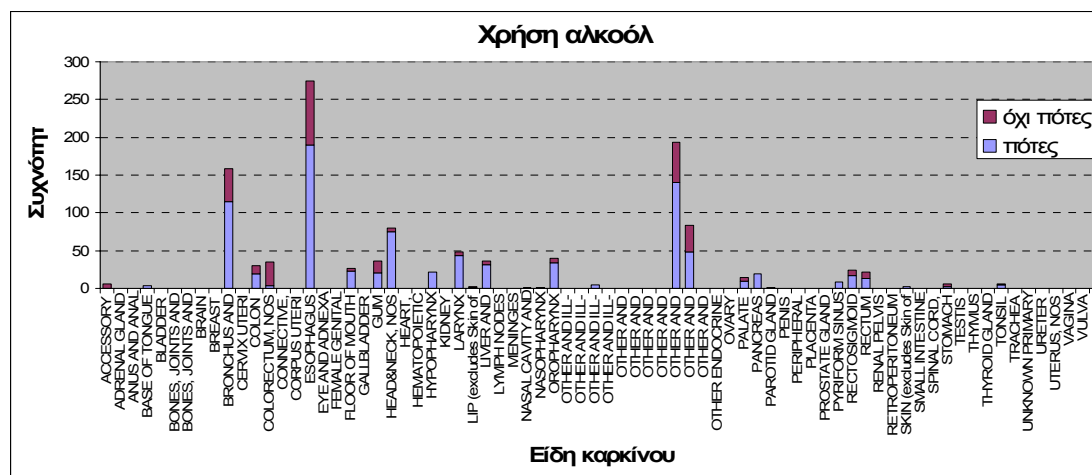
**Σχήμα 42. Συχνότητες νεοπλασμάτων σε σχέση με τους καπνιστές και μη καπνιστές.**

Φαίνεται ξεκάθαρα ο κίνδυνος του καπνίσματος για την εμφάνιση καρκίνου του πνεύμονα, της κύστεως, του οισοφάγου και του στόματος.

Είναι αξιοσημείωτο, ότι πέρα από τα κλασικά σημεία που επηρεάζουν οι τοξίνες του καπνού (πνεύμονας, οισοφάγος, στοματική κοιλότητα), και η ουροδόχος κύστη επηρεάζεται, πιθανόν μέσω της αποβολής τοξινών από τα ούρα. Οι τοξίνες του καπνού φαίνεται ότι αθροίζονται στην κύστη πριν από την ούρηση και η τοπική αύξηση της συγκέντρωσής τους προκαλεί αύξηση των νεοπλασμάτων του οργάνου.

Στον παρακάτω πίνακα υπολογίζεται ο κίνδυνος εμφάνισης καρκίνου των καπνιστών σε σχέση με τους μη καπνιστές. Ο κίνδυνος υπολογίζεται ως κλάσμα= καπνιστές/(καπνιστές+μη καπνιστές).

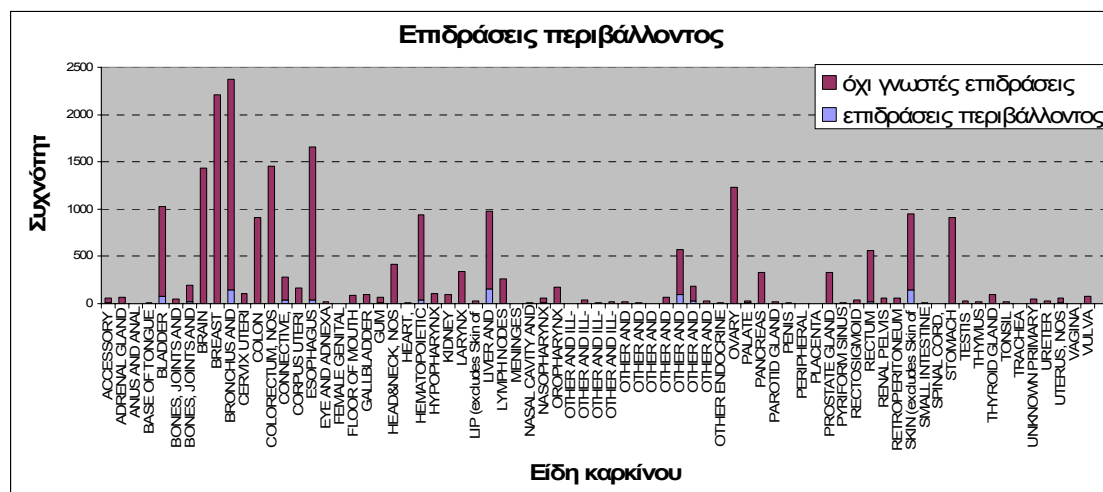
Είδος καρκίνου	Συχνότητα εμφάνισης		Κίνδυνος εμφάνισης του καρκίνου σε καπνιστές
	καπνιστές	μη καπνιστές	
RECTUM	5	17	23%
STOMACH	2	5	29%
COLORECTUM, NOS	28	43	39%
BREAST	7	10	41%
COLON	13	17	43%
ACCESSORY SINUSES	3	3	50%
GUM	25	20	56%
RECTOSIGMOID JUNCTION	15	9	63%
RENAL PELVIS	19	11	63%
OTHER AND UNSPECIFIED PARTS OF TONGUE	78	35	69%
OTHER AND ILL-DEFINED SITES WITHIN RESPIRATORY SYSTEM AND INTRATHORACIC ORGANS	13	5	72%
FLOOR OF MOUTH	40	14	74%
PANCREAS	20	7	74%
BLADDER	224	75	75%
ESOPHAGUS	285	89	76%
BRONCHUS AND LUNG	883	240	79%
BASE OF TONGUE	5	1	83%
OTHER AND UNSPECIFIED PARTS OF MOUTH	180	36	83%
URETER	5	1	83%
SKIN (excludes Skin of vulva, Skin of penis, Skin of scrotum )	28	5	85%
PYRIFORM SINUS	7	1	88%
HEAD&NECK, NOS	77	10	89%
LIVER AND INTRAHEPATIC BILE DUCTS	8	1	89%
PALATE	18	1	95%
OROPHARYNX	38	2	95%
HYPOPHARYNX	21	1	95%
LARYNX	80	3	96%
BONES, JOINTS AND ARTICULAR CARTILAGE OF OTHER AND UNSPECIFIED SITES	1	0	100%
LIP (excludes Skin of lip )	2	0	100%
NASAL CAVITY AND MIDDLE EAR	1	0	100%
NASOPHARYNX	1	0	100%
OTHER AND ILL-DEFINED SITES IN LIP, ORAL CAVITY AND PHARYNX	5	0	100%
PAROTID GLAND	1	0	100%
TONSIL	8	0	100%



**Σχήμα 43. Συχνότητες νεοπλασμάτων σε σχέση με την κατανάλωση οινοπνεύματος.**

Στον παρακάτω πίνακα φαίνεται ο κίνδυνος εμφάνισης καρκίνου στους πότες σε σχέση με αυτούς που δεν είναι πότες. Ο κίνδυνος υπολογίζεται ως κλάσμα= πότες/(πότες+όχι πότες).

Είδος καρκίνου	Συχνότητα		Κίνδυνος εμφάνισης καρκίνου στους πότες
	πότες	όχι πότες	
ACCESSORY SINUSES	0	6	0%
COLORECTUM, NOS	4	31	11%
LIP (excludes Skin of lip )	1	1	50%
STOMACH	3	3	50%
OTHER AND UNSPECIFIED PARTS OF TONGUE	48	36	57%
GUM	21	15	58%
RECTUM	13	9	59%
COLON	19	11	63%
PALATE	10	5	67%
SKIN (excludes Skin of vulva, Skin of penis, Skin of scrotum )	2	1	67%
ESOPHAGUS	190	85	69%
RECTOSIGMOID JUNCTION	17	7	71%
OTHER AND UNSPECIFIED PARTS OF MOUTH	140	54	72%
BRONCHUS AND LUNG	115	43	73%
TONSIL	5	1	83%
OROPHARYNX	34	6	85%
FLOOR OF MOUTH	23	4	85%
LIVER AND INTRAHEPATIC BILE DUCTS	32	4	89%
LARYNX	44	4	92%
HEAD&NECK, NOS	75	5	94%
BASE OF TONGUE	4	0	100%
HYPOPHARYNX	22	0	100%
NASAL CAVITY AND MIDDLE EAR	1	0	100%
NASOPHARYNX	1	0	100%
OTHER AND ILL-DEFINED SITES IN LIP, ORAL CAVITY AND PHARYNX	5	0	100%
PANCREAS	19	0	100%
PAROTID GLAND	1	0	100%
PYRIFORM SINUS	8	0	100%



**Σχήμα 44. Συνολικές καταγραφές στη βάση δεδομένων σε σχέση με τις επιδράσεις του περιβάλλοντος.**

Όπως φαίνεται στη βάση δεδομένων η πλειοψηφία των καταγραφών δεν αναφέρει περιβαλλοντικές επιδράσεις. Οι επιδράσεις του περιβάλλοντος καταγράφονται κυρίως στους καρκίνους της ουροδόχου κύστεως, του πνεύμονα, του ήπατος και του δέρματος.

#### IV. ΣΥΖΗΤΗΣΗ

Το γονίδιο *p53* έχει ερευνηθεί εκτεταμένως διότι έχει αναγνωριστεί η μεγάλη του σημασία στη ρύθμιση των λειτουργιών των κυττάρων. Έχει γίνει μεγάλη προσπάθεια αποκάλυψης των μεταλλάξεων που υπόκειται στα καρκινικά κύτταρα και έχει καταγραφεί το φάσμα των μεταλλάξεων του και στις γενετικές μεταλλάξεις και στις σωματικές. Αυτές οι βάσεις προσφέρονται για ανάλυση διότι περιέχουν πολλές λεπτομερείς καταγραφές.

*Η ανάλυση των παραπάνω βάσεων προσφέρει σημαντικές ενδείξεις για το βαθμό μεθυλίωσης του γονιδίου.* Θεωρούμε ότι οι μεταλλάξεις G>A και C>T προέρχονται από μεθυλίωση της κυτοσίνης. Παρατηρούμε ότι το ποσοστό των μεταλλάξεων G:C>A:T at CpG από μεθυλίωση είναι μεγαλύτερο στα γεννητικά κύτταρα, ενώ αντίθετα το ποσοστό των μεταλλάξεων G:C>A:T που δεν βρίσκονται σε θέσεις CpG είναι μεγαλύτερο στα σωματικά κύτταρα. Αυτό είναι ενδεικτικό ότι τα στα σωματικά κύτταρα συμβαίνει και μεθυλίωση της κυτοσίνης εκτός θέσεων CpG.

Το διπλάσιο ποσοστό μεταλλάξεων G:C>A:T at CpG στις γενετικές μεταλλάξεις σε σύγκριση με τις σωματικές, προκύπτει επειδή στις γενετικές μεταλλάξεις επικρατούν οι μεταλλάξεις των hotspots. Αντίθετα, στα σωματικά κύτταρα παρατηρούμε εντονότερη μεθυλίωση των non-CpG θέσεων κυτοσίνης και επακόλουθη αύξηση των μεταλλάξεων σ' αυτές. Επίσης, η αύξηση στις μεταλλάξεις G:C>T:A θα μπορούσε να οφείλεται σε μεθυλίωση της γουανίνης από τοξικούς παράγοντες.[62] Η αύξηση των μεταλλάξεων A:T>C:G δείχνει τη γενικότερη τάση για αύξηση των μεταλλάξεων που δεν οφείλονται σε μεθυλίωση και γενετική αστάθεια.

Στις περιπτώσεις που τα hotspots εντοπίζονται μόνο σε μία από τις δύο βάσεις δεδομένων και αφορούν τον ίδιο ιστό, μπορούμε να υποθέσουμε ότι:

(α) δεν πραγματοποιούνται μεταλλάξεις σε αυτές τις θέσεις στην άλλη σειρά, (β) οι μεταλλάξεις σ' αυτές τις θέσεις δεν προκαλεί καρκινογένεση λόγω διαφορετικών συνθηκών στο μικροπεριβάλλον του πυρήνα, (γ) δεν είναι εφικτή η αποτελεσματική επιδιόρθωση ή (δ) πρωτεΐνες που προσκολλώνται στην έλικα του DNA δεν επιτρέπουν τη μετάλλαξη σε αυτή τη θέση.

*Αξιολόγηση των προγραμμάτων ανάλυσης που χρησιμοποιήθηκαν στην εργασία.* Το iMARS παρατηρούμε ότι παρέχει χρήσιμες γενικές πληροφορίες για την αλληλουχία του γονιδίου και τις μεταλλάξεις του. Επειδή οι αναλύσεις του είναι τυποποιημένες δεν είναι εφικτό να λάβει υπόψη του πάντα τις σωστές μεταβλητές ώστε να καταλήξει σε στατιστικώς σημαντικά συμπεράσματα. Για παράδειγμα, όσον αφορά



την εκτίμηση της προέλευσης των μεταλλάξεων του τύπου G:C>A:T, το πρόγραμμα iMARS δεν επιτρέπει το διαχωρισμό σε G:C>A:T at CpG και G:C>A:T non CpG, με αποτέλεσμα να μην είναι δυνατή η ανάλυση που αναφέρθηκε παραπάνω. Το πρόγραμμα διαχείρισης spreadsheet παρέχει μεγάλη ελευθερία ανάλυσης των δεδομένων και πολλές επιλογές σχηματοποίησης τους. Επομένως οι αναλύσεις μας βασίστηκαν κυρίως στο πρόγραμμα spreadsheet ώστε να ληφθούν υπόψη οι επιθυμητές μεταβλητές και παράμετροι.

*Σταθμισμένη μεταλλαξιμότητα.* Για να υπολογιστεί ο ρυθμός μετάλλαξης ενός κωδικίου διαιρείται ο συνολικός αριθμός μεταλλάξεων του κωδικίου προς τον συνολικό αριθμό που εμφανίζεται το κωδικίο στο γονίδιο. Το πηλίκο που προκύπτει ονομάζεται σταθμισμένη μεταλλαξιμότητα. Ειδικά για ορισμένα κωδικία που έχουν υψηλή σταθμισμένη μεταλλαξιμότητα (CCG, CGG, TCC, TGC, TTT) μπορούμε να υποθέσουμε ότι οφείλεται στην περιεκτικότητά τους σε μεθυλιωμένες κυτοσίνες και για την τριπλέτα TTT ότι αποτελεί στοιχείο αποελίκωσης της αλυσίδας του DNA.

*Γενετικός κώδικας.* Όπως γνωρίζουμε, ο γενετικός κώδικας είναι εκφυλισμένος. Αυτό σημαίνει ότι αμινοξέα που αντιπροσωπεύονται με ένα ή δύο κωδικία είναι πιο επιρρεπή σε μεταλλάξεις από αυτά που αντιπροσωπεύονται από μεγαλύτερο αριθμό κωδικίων. Από τα σχήματα προκύπτει ότι από κάθε αμινοξύ και κατ' επέκταση από τα κωδικία που το εκπροσωπούν προκύπτουν συγκεκριμένα είδη μεταλλάξεων.

Είναι αξιοσημείωτο, ότι η αργινίνη παρόλο που εκπροσωπείται από 6 κωδικία, το μέγιστο αριθμό δηλαδή, μεταλλάσσεται με την μέγιστη συχνότητα. Έχει την ιδιότητα ως βασικό αμινοξύ να συνδέεται με το DNA και επομένως η σημειακή απώλειά της μπορεί να στερεί μεγάλο μέρος της λειτουργικότητας του p53. Παρατηρούμε ότι κυριαρχούν οι μεταλλάξεις μεθυσίωσης G:C>A:T at CpG με 116 εγγραφές στο σύνολο των 132 γενετικών μεταλλάξεων της αργινίνης ανά οικογένεια και με 5.900 καταγραφές στο σύνολο των 6.840 σωματικών μεταλλάξεων της αργινίνης. Τα κωδικία της αργινίνης που ξεκινούν με την τριπλέτα CG μεταλλάσσονται πολύ περισσότερο από τα κωδικία της αργινίνης που ξεκινούν με την τριπλέτα AG. Με τη δοκιμασία  $\chi^2$  συμπεραίνουμε ότι αυτές οι υποθέσεις έχουν  $p=0,0000$  για τις σωματικές μεταλλάξεις και  $p=0,0000$  για τις γενετικές μεταλλάξεις.

Από τις αρχές της δημιουργίας της ζωής ο γενετικός κώδικας θα προσπαθούσε να εξαλείψει την αλλαγή του μηνύματος που επιφέρουν οι μεταλλάξεις. Έτσι, για παράδειγμα, θα εκπροσωπούσε υποθετικά την αργινίνη με κωδικία να αρχίζουν από κυτοσίνη και άλλα από θυμίνη, ώστε να μην επηρεαστεί το μήνυμα αν μεθυλιωθεί

η κυτοσίνη και μετατραπεί σε θυμίνη. Αυτό όμως δεν γίνεται. Αντίθετα, βλέπουμε ότι ο γενετικός κώδικας εξαλείφει την επίδραση μετάλλαξης της τριπλέτας μετά από τη θέση CG, καθώς όποιο και να είναι το τρίτο αμινοξύ κωδικοποιείται η αργινίνη. Αυτό είναι αξιοσημείωτο χωρίς προφανή εξήγηση.

Επίσης, η προλίνη, γλυκίνη, βαλίνη εκπροσωπούνται από 4 κωδίκια και μεταλλάσσονται πολύ έντονα. Το ίδιο και η κυστεΐνη, τυροσίνη και γλουταμικό οξύ που εκπροσωπούνται από 2 κωδίκια. Αντίθετα, το αμινοξύ φαινυλαλανίνη, που εκπροσωπείται από 2 κωδίκια, και το αμινοξύ τρυπτοφάνη, που εκπροσωπείται από 1 κωδικίο, παρουσιάζουν το μικρότερο ποσοστό μεταλλάξεων.

Άρα τα κωδίκια της αργινίνης, προλίνης, γλυκίνης, βαλίνης, κυστεΐνης, τυροσίνης και γλουταμικού οξέος μεταλλάσσονται εύκολα. Η τριπλέτα TTC μεταλλάσσεται συνήθως σε TTT. Η TGG μεταλλάσσεται πολύ δύσκολα. Πολύ δύσκολα μεταλλάσσεται και η τριπλέτα TAC, όπως προέκυψε από τη χαμηλή σταθμισμένη μεταλλαξιμότητά της. Αν προκύψει μετάλλαξη μεθυλίωσης της κυτοσίνης C>T (TAC > TAT), είναι σιωπηρή και κλινικώς μη παρατηρήσιμη. Άρα οι κλινικώς παρατηρήσιμες μεταλλάξεις της τυροσίνης οφείλονται κυρίως σε μεταλλάξεις της τριπλέτας TAT.

Εκτός από τις θέσεις «κλειδιά» CpG, οι περισσότερες από τις οποίες κωδικοποιούν την αργινίνη και είναι μεθυλιωμένες, σε πληθώρα άλλων θέσεων ενεργοποιείται η μεταλλαξιγένεση σε σποραδικό καρκίνο, ιδίως σε κωδίκια που δίνουν G:C>A:T μεταλλάξεις και επομένως θα μπορούσαν να είναι μεθυλιωμένα, αλλά μόνο με διαδικασία ασύμμετρης μεθυλίωσης.

*Η συμβολή των CpG θέσεων στη μεταλλαξιγένεση.* Όπως προκύπτει από τα αποτελέσματα της εργασίας, σε πολλές γενετικές θέσεις CpG, παρόλο που θεωρείται ότι είναι μεθυλιωμένη η κυτοσίνη, δεν υπερिशύουν οι μεταλλάξεις μεθυλίωσης (G:C>A:T at CpG) και στα γενετικά και στα σωματικά κύτταρα. Οι θέσεις αυτές εντοπίζονται στα κωδίκια 153, 156, 157, 158, 170, 181, 202, 267 και 290. Επιπλέον, στα κωδίκια 196 και 306 υπερिशύουν οι μεταλλάξεις μεθυλίωσης G:C>A:T at CpG, αλλά οι θέσεις αυτές δεν είναι hotspots ούτε στις γενετικές ούτε στις σωματικές μεταλλάξεις. Αυτό σημαίνει ότι η θεωρία του Pfeifer [51] για τις μεταλλάξεις της κυτοσίνης σε θυμίνη λόγω μεθυλίωσης και την πρόκληση μεγάλου αριθμού μεταλλάξεων έχει κενά. Δεν αποτελεί μοναδική και απαραίτητη προϋπόθεση για υψηλό ποσοστό μεταλλαξιγένεσης η μεθυλίωση της κυτοσίνης του δινουκλεοτιδίου CpG.

*Μέσος όρος ηλικίας εμφάνισης καρκίνου των σιωπηρών σωματικών μεταλλάξεων.* Παρατηρούμε ότι και οι σιωπηρές σωματικές

μεταλλάξεις έχουν τον ίδιο μέσο όρο ηλικίας εμφάνισης καρκίνου με τις υπόλοιπες κατηγορίες μεταλλάξεων. Από τη βάση δεδομένων προκύπτει ότι η πλειονότητα των σιωπηρών μεταλλάξεων είναι πρωτογενής και δεν συνοδεύεται από άλλες μεταλλάξεις. Επίσης διαπιστώσαμε ότι κατανέμονται σε ολόκληρη την αλληλουχία του γονιδίου και δεν είναι εστιασμένες. Για να εξηγήσουμε αυτό το φαινόμενο μπορούμε να υποθέσουμε ότι λόγω της σιωπηρής μετάλλαξης «χάνουν» τα ένζυμα του ματίσματος (splicing) τη σωστή θέση και να προκύπτει μεγαλύτερο ή και μικρότερο μεταγράφημα.

*Οι γενετικές μεταλλάξεις των ενδονίων.* Οι γενετικές μεταλλάξεις των ενδονίων (intronic) δεν επηρεάζουν άμεσα την παραγόμενη πρωτεΐνη, αλλά έχουν μέσο όρο ηλικίας εμφάνισης καρκίνου που αγγίζει τα 50 έτη, ενώ ο μέσος όρος είναι 67 έτη. Είναι πιθανό οι μεταλλάξεις αυτές να αλλοιώνουν θέσεις ματίσματος και να δίνουν τελικά διαφορετικό μέγεθος μεταγραφήματος από αυτό που προκύπτει φυσιολογικά. Είναι πιθανό σε εκείνο το διάστημα της ζωής να απαιτείται η παραγωγή μιας ισομορφής του ενζύμου, αλλά οι θέσεις ματίσμάτος του (splice sites) έχουν αλλοιωθεί από τις μεταλλάξεις των ενδονίων. Αυτή η θεωρία ενισχύεται από το γεγονός ότι οι πλειοψηφία των καταγραφών ανήκει στην πρώτη γενεά που είναι τα άτομα με τη μεγαλύτερη ηλικία. Άρα μπορούμε να υποθέσουμε ότι οι μεταλλάξεις των ενδονίων πιθανόν να σχετίζονται με μεταγραφικές τροποποιήσεις και εναλλακτικούς τρόπους ματίσματος που συμβαίνουν σε ορισμένα ηλικιακά στάδια της ζωής του ανθρώπου.

*Πιθανοί μηχανισμοί που προάγουν ή αποτρέπουν τη μεταλλαξιγένεση.* Είναι αξιοσημείωτο ότι σε αρκετά κωδικία και στις γενετικές και στις σωματικές μεταλλάξεις μεταλλάσσεται περισσότερο η γουανίνη και λιγότερο η κυτοσίνη. Το συμπέρασμα αυτό στηρίζεται και από την ανάλυση του iMARS, όπου φαίνεται ότι οι περισσότερες μεταλλάξεις της κωδικής αλυσίδας στο ζεύγος γουανίνη-κυτοσίνη γίνονται από την γουανίνη. Ένας πιθανός μηχανισμός είναι να υποθέσουμε ότι μεθυλιώνεται η κυτοσίνη της συμπληρωματικής αλυσίδας απέναντι από τη γουανίνη και προκαλεί ισχυρή μεταλλαξιγόνο δράση.

Στις σωματικές μεταλλάξεις παρατηρούμε ότι στο κωδικίο 282 (CGG), που αποτελεί hotspot, υπάρχουν 476 καταγραφές C>T μεταβάσεων στο σύνολο των 586 μεταλλάξεων. Αυτό το φαινόμενο προάγεται σε σημαντικό βαθμό από τη μεθυλίωση της κυτοσίνης της κωδικής αλυσίδας. Αντίθετα, στο διπλανό κωδικίο 283 (CGC) παρατηρείται πολύ μικρότερος αριθμός μεταλλάξεων, και οι καταγραφές C>T είναι 27 στο σύνολο των 95. Και τα δύο κωδικία κωδικοποιούν το αμινοξύ αργινίνη. Οι μεταλλάξεις C>T του κωδικίου 282 οδηγούν στο αμινοξύ τρυπτοφάνη και του 283 στο αμινοξύ

κυστεΐνη. Αυτό σημαίνει ότι πιθανόν να υπάρχει κάποιος παράγοντας που αποτρέπει τη μετάλλαξη του κωδικίου 283 στη θέση CpG, είτε ότι η μετάλλαξη του κωδικίου 283 δεν είναι τόσο τοξική όσο του 282.

*Διεισδυτικότητα των γενετικών μεταλλάξεων.* Σε μια ορισμένη ομάδα κωδικίων των γενετικών μεταλλάξεων οι φορείς εμφανίζουν τη νόσο σε μεγάλη ηλικία (κωδικία 196, 234, 245, 248, 282). Στην περίπτωση αυτή, μπορεί να υποθεθεί ότι οι μεταλλάξεις είτε είναι μικρής σημασίας για τη λειτουργία της πρωτεΐνης, είτε σχετίζονται με αύξηση της γενετικής αστάθειας του γονιδίου.

*Μεταλλάξεις δίχως οικογενειακό ιστορικό.* Οι πιο επιθετικές γενετικές μεταλλάξεις είναι οι μεταλλάξεις δίχως οικογενειακό ιστορικό (No Family History) επειδή έχουν το μικρότερο προσδόκιμο επιβίωσης. Αφορούν κυρίως τα κωδικία 273 και 278, αλλά και 151, 158, 161, 167, 189, 241, 242, 245, 248, 273, 282 και 293. Εντοπίζονται σε περιοχές πλούσιες σε G:C περιεχόμενο και από την ανάλυση της αλληλουχίας τους απορρέει το συμπέρασμα ότι δεν περιλαμβάνουν το δινουκλεοτίδιο CpG εκτός από την 158. Άρα οι μεταλλάξεις χωρίς ιστορικό ξεκινούν, κατά πλειοψηφία, από αλυσίδες πλούσιες σε G:C και όχι από δινουκλεοτίδια CpG.

Είναι αξιοσημείωτο το μεγάλο ποσοστό των μεταλλάξεων χωρίς ιστορικό στον καρκίνο των επινεφριδίων και το μηδενικό ποσοστό στον καρκίνο του στομάχου. Το γεγονός αυτό υποδηλώνει ότι ο καρκίνος των επινεφριδίων σε μεγάλο ποσοστό προέρχεται από οικογένειες χωρίς οικογενειακό ιστορικό. Αντίθετα, καρκίνο του στομάχου είναι συντριπτικά πιθανότερο να πάθουν άτομα με οικογενειακό ιστορικό.

*Γενικά συμπεράσματα.* Είναι σαφές ότι ακόμη δεν είναι εφικτή η ερμηνεία όλων των ιδιομορφιών του φάσματος των γενετικών μεταλλάξεων. Η ερμηνεία με βάση την γενετική αλληλουχία και τη θεωρία της μεθυλίωσης αποτελούν μόνο ένα μικρό κομμάτι του παζλ. Είναι πιθανή η εμπλοκή και άλλων μηχανισμών μεθυλίωσης, οι οποίοι να σχετίζονται με τη μεταγραφική ρύθμιση του *p53* σε κάποια ηλικιακά στάδια του ανθρώπου. Το μοντέλο της επιγενετικής τροποποίησης φαίνεται ότι είναι σύνθετο, αφορά και τις δύο έλικες του DNA και μεταβάλλεται ανάλογα με το είδος του καρκίνου. Οι αναλύσεις του φάσματος των μεταλλάξεων της παρούσας εργασίας είναι πολύτιμα εργαλεία για τη διατύπωση θεωρητικών μοντέλων, τα οποία στη συνέχεια μπορούν να διερευνηθούν πειραματικά.

## V. ΒΙΒΛΙΟΓΡΑΦΙΑ

[1] Rakesh Singal and Gordon D. Ginder. DNA Methylation. *Blood*, Vol 93, No 12 (June 15), 1999: pp 4059-4070

[2] Bird AP: Gene number, noise reduction and biological complexity. *Trends Genet* 11:94, 1995

[3] Weissbach A: A chronicle of DNA methylation (1948-1975). *EXS* 64:1, 1993

[4] Russo, V.E.A., Martienssen, R.A., and Riggs, A.D. 1996. *Epigenetic mechanisms of gene regulation*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

[5] Bird A. DNA methylation patterns and epigenetic memory. *Genes Dev* 16; 6-21 (2002)

[6] Gonzalgo ML, Liang G, Spruck CHr, Zingg JM, Rideout WMr, Jones PA: Identification and characterization of differentially methylated regions of genomic DNA by methylation-sensitive arbitrarily primed PCR. *Cancer Res* 57:594, 1997

[7] Bender CM, Zingg JM, Jones PA: DNA methylation as a target for drug design. *Pharmaceutical Res* 15:175, 1998

[8] AM Jubb, SM Bell, P Quirke. Methylation and colorectal cancer. *J Pathol* 2001; 195: 111-134.

[9] A.D. Riggs, P.A. Jones, 5-Methylcytosine, gene regulation, and cancer, *Adv. Cancer Res.* 40\_1983.1-30.

[10] A.P. Bird, The essentials of DNA methylation, *Cell* 70 \_1992.5-8.

[11] L.A. Frederico, T.A. Kunkel, B.R. Shaw, A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy, *Biochemistry* 29\_1990.2532-2537.

[12] T. Lindahl, Instability and decay of the primary structure of DNA, *Nature* 362\_1993.709-715.

[13] J.C. Shen, W.M. Rideout III, P.A. Jones, The rate of hydrolytic deamination of 5-methylcytosine in double-stranded

DNA, *Nucleic Acids Res.* 22\_1994.972–976.

[14] R.Y. Wang, K.C. Kuo, C.W. Gehrke, L.H. Huang, M. Ehrlich, Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA, *Biochim. Biophys. Acta* 697\_1982.371–377.

[15] M. Ehrlich, K.F. Norris, R.Y. Wang, K.C. Kuo, C.W. Gehrke, DNA cytosine methylation and heat-induced deamination, *Biosci. Rep.* 6\_1986.387–393.

[16] M. Ehrlich, X.-Y. Zhang, N.M. Inamdar, Spontaneous deamination of cytosine and 5-methylcytosine residues in DNA and replacement of 5-methylcytosine residues with cytosine residues, *Mutat. Res.* 238\_1990.277–286.

[17] D.N. Cooper, H. Youssoufian, The CpG dinucleotide and human genetic disease, *Hum. Genet.* 78\_1988.151–155.

[18] C. Schmutte, A.S. Yang, R.W. Beart, P.A. Jones, Base excision repair of U:G mismatches at a mutational hotspot in the *p53* gene is more efficient than base excision repair of T:G mismatches in extracts of human colon tumors, *Cancer Res.* 55\_1995.3742–3746.

[19] P. Neddermann, P. Gallinari, T. Lettieri, D. Schmid, O. Truong, J.J. Hsuan, K. Wiebauer, J. Jiricny, Cloning and expression of human GrT mismatch-specific thymine-DNA glycosylase, *J. Biol. Chem.* 271\_1996.12767–12774.

[20] M. Sapparbaev, J. Laval, 3,*N*4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for *Escherichia coli* double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase, *Proc. Natl. Acad. Sci. USA* 95\_1998.8508–8513.

[21] B. Hang, M. Medina, H. Fraenkel-Conrat, B. Singer, A 55-kDa protein isolated from human cells shows DNA glycosylase activity toward 3,*N*4-ethenocytosine and the GrT mismatch, *Proc. Natl. Acad. Sci. USA* 95\_1998.13561–13566.

[22] A.S. Yang, M.L. Gonzalgo, J.M. Zingg, R.P. Millar, J.D. Buckley, P.A. Jones, The rate of CpG mutation in Alu repetitive elements within the *p53* tumor suppressor gene in the primate germline, *J. Mol. Biol.* 258\_1996.240–250.

- [23] A. Riccio, L.A. Aaltonen, A.K. Godwin, A. Loukola, A. Percesepe, R. Salovaara, V. Masciullo, M. Genuardi, M. Paravatou-Petsotas, D.E. Bassi, B.A. Ruggeri, A.J. Klein-Szanto, J.R. Testa, G. Neri, A. Bellacosa, The DNA repair gene MBD4\_MED1 is mutated in human carcinomas with microsatellite instability, *Nat. Genet.* 23\_1999.266–268.
- [24] S.S. Sommer, Recent human germ-line mutation: inferences from patients with hemophilia B, *Trends Genet.* 11\_1995. 141–147.
- [25] C. Schmutte, P.A. Jones, Involvement of DNA methylation in human carcinogenesis, *Biol. Chem.* 379\_1998.377–388.
- [26] J. Wagner, R.C.C. Hu, B.N. Ames, Endogenous oxidative damage of deoxycytidine in DNA, *Proc. Natl. Acad. Sci. USA* 89\_1992.3380–3384.
- [27] E. Eisenstadt, A.J. Warren, J. Porter, D. Atkins, J.H. Miller, Carcinogenic epoxides of benzo\_ *a* .pyrene and cyclopenta\_ *cd* .pyrene induce base substitutions via specific transversions, *Proc. Natl. Acad. Sci. USA* 79\_1982.1945–1949.
- [28] B.L. Van Duuren, The polynuclear aromatic hydrocarbons in cigarette smoke condensate II, *J. Natl. Cancer Inst.* 21\_1958. 623–630.
- [29] E.L. Wynder, D. Hoffmann, A study of tobacco carcinogenesis: VII. The role of higher polycyclic hydrocarbons, *Cancer* 12\_1959.1079–1086.
- [30] M.F. Denissenko, A. Pao, G.P. Pfeifer, M.-S. Tang, Slow repair of bulky DNA adducts along the nontranscribed strand of the human *p53* gene may explain the strand bias of transversion mutations in cancers, *Oncogene* 16 \_1998. 1241–1247
- [31] Schmid CW: Alu: Structure, origin, evolution, significance and function of one-tenth of human DNA. *Prog Nucleic Acid Res Mol Biol* 53:283, 1996
- [32] Doerfler W, Schubbert R, Heller H, Kammer C, Hilger-Eversheim K, Knoblauch M, Remus R: Integration of foreign DNA and its consequences in mammalian systems. *Trends Biotechnol* 15:297, 1997

- [33] Barletta J, Greer SB: Methylation of HSV-1 DNA as a mechanism of viral inhibition: Studies of an analogue of methyldeoxycytidine: Trifluoromethyldeoxycytidine (F3mdCyd). *Antiviral Res* 18:1,1992
- [34] S. Tommasi, M.F. Denissenko, G.P. Pfeifer, Sunlight induces pyrimidine dimers preferentially at 5-methylcytosine bases, *Cancer Res.* 57\_1997.4727–4730.
- [35] Robertson K. DNA methylation and chromatin – unraveling the tangled web. *Oncogene* 21; 5361-5379 (2002)
- [36] Hendrich B, Tweedie S. The methyl-CpG binding domain and the evolving role of DNA methylation in animals. *Trends Genet* 19; 269-277 (2003)
- [37] Pradhan S, Esteve PO. Mammalian DNA (cytosine-5) methyltransferases and their expression. *Clin Immunol* 2003;109:6–16.
- [38] Hermann A, Gowher H, Jeltsch A. Biochemistry and biology of mammalian DNA methyltransferases. *Cell Mol Life Sci* 2004;61:2571–87.
- [39] Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *JBiol Chem* 1999;274:33002–10.
- [40] Ramsahoye BH, Biniszkiwicz D, Lyko F, Clark V, Bird AP, Jaenisch R. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci USA* 2000;97:5237–42.
- [41] Hsieh CL. In vivo activity of murine de novo methyltransferases Dnmt3a and Dnmt3b. *Mol Cell Biol* 1999;19:8211–8.
- [42] Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69; 915-926 (1992)
- [43] Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99; 247-257 (1999)



- [44] Brandeis, M., Frank, D., Keshet, I., Siegried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. 1994. Sp1 elements protect a CpG island from de novo methylation. *Nature* 371: 435–438.
- [45] MacLeod, D., Clark, V., and Bird, A. 1999. Absence of genomewide changes in DNA methylation during development of the zebrafish (*Danio rerio*). *Nat. Genet.* 23: 139–140.
- [46] Hertz, J.M., Schell, G., and Doerfler, W. 1999. Factors affecting de novo methylation of foreign DNA in mouse embryonic stem cells. *J. Biol. Chem.* 274: 24232–24240.
- [47] [http://p53.free.fr/Database/p53\\_mutation.html](http://p53.free.fr/Database/p53_mutation.html)
- [48] Pfeifer, G.P., and Holmquist, G.P. 1997. Mutagenesis in the *P53* gene. *Biochim. Biophys. Acta* 1333: 1-8.
- [49] Soussi, T. 2000. The *p53* tumor suppressor gene: from molecular biology to clinical investigation. *Ann. N.Y. Acad. Sci.* 910: 121-137.
- [50] Soussi, T., and Beroud, C. 2003. Significance of TP53 mutations in human cancer: a critical analysis of mutations at CpG dinucleotides. *Hum. Mutat.* 21: 192-200.
- [51] Pfeifer, G.P. 2000. p53 mutational spectra and the role of methylated CpG sequences. *Mutat. Res.* 450: 155-166.
- [52] M.Gasco, I.G.Yulug, T.Crook. TP53 mutations in familial breast cancer: functional aspects, *Hum. Mutat.* 21 (2003) 301-306.
- [53] S. Strano, M. Rossi, G. Fontemaggi, E. Munarriz, S. Soddu, A. Sacchi, G. Blandino, From p63 to p53 across p73, *FEBS Lett.* 490 (2001) 163– 170.
- [54] J. Benard, S. Douc-Rasy, J.C. Ahomadegbe, TP53 family members and human cancers, *Hum. Mutat.* 21 (2003) 182– 191.
- [55] H. Schmale, C. Bamberger, A novel protein with strong homology to the tumor suppressor p53, *Oncogene* 15 (1997) 1363–1367.
- [56] M. Kaghad, H. Bonnet, A. Yang, L. Creancier, J.C. Biscan, A. Valent, A. Minty, P. Chalon, J.M. Lelias, X. Dumont, P. Ferrara, F. McKeon, D. Caput, Monoallelically expressed gene related to p53 at

1p36, a region frequently deleted in neuroblastoma and other human cancers, *Cell* 90 (1997) 809– 819.

[57] M.C. Marin, W.G. Kaelin Jr., p63 and p73: old members of a new family, *Biochim. Biophys. Acta* 1470 (2000) M93–M100.

[58] M.A. Lohrum, K.H. Vousden, Regulation and function of the p53-related proteins: same family, different rules, *Trends Cell. Biol.* 10 (2000) 197–202.

[59] Pfeifer, G.P., Denissenko, M.F., Olivier, M., Tretyakova, N., Hecht, S.S., and Hainaut, P. 2002. Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* 21: 7435-51.

[60] Olivier, M., Eeles, R., Hollstein, M., Khan, M.A., Harris, C.C., and Hainaut, P. 2002. The IARC TP53 Database: new online mutation analysis and recommendations to users. *Hum. Mutat.* 19: 607-614.

[61] S.Tornaletti, G.P.Pfeifer. Complete and tissue-independent methylation of CpG sites in the p53 gene: implications for mutations in human cancers, *Oncogene* 10 (1995) 1493-1499.

[62] Yoon, J.H., Smith, L.E., Feng, Z., Tang, M., Lee, C.S., and Pfeifer, G.P. 2001. Methylated CpG dinucleotides are the preferential targets for G-to-T transversion mutations induced by benzo[a]pyrene diol epoxide in mammalian cells: similarities with the p53 mutation spectrum in smoking-associated lung cancers. *Cancer Res.* 61: 7110-7117.

[63] Genetic instability of p53 in sporadic and BRCA1/2-associated breast and ovarian cancers. S koudou, A Malousi, A Kyventidis, K Fragou, N Maglaveras (submitted for publication)

[64] L.L.Hsieh, P.F.Wang, I.H.Chen, C.T.Liao, H.M.Wang, M.C.Chen, J.T.Chang, A.J.Cheng. Characteristics of mutations in the p53 gene in oral squamous cell carcinoma associated with betel quid chewing and cigarette smoking in Taiwanese, *Carcinogenesis* 22 (2001) 1497-1503.

[65] X.Agirre, J.L.Vizmanos, M.J.Calasan, M.Garcia-Delgado, M.J.Larrayoz, F.J.Novo. Methylation of CpG dinucleotides and/or CCWGG motifs at the promoter of TP53 correlates with decreased gene expression in a subset of acute lymphoblastic leukemia patients, *Oncogene* 22 (2003) 1070-1072.

- [66] E.M.LeProust, C.E.Pearson, R.R.Sinden, X.Gao. Unexpected formation of parallel duplex in GAA and TTC trinucleotide repeats of Friedreich's ataxia, *J. Mol. Biol.* 302 (2000) 1063-1080.
- [67] Hsieh, L.L., Wangm P.F., Chenm I.H., Liao, C.T., Wang, H.M., Chen, M.C., Chang, J.T., and Cheng, A.J. 2001. Characteristics of mutations in the p53 gene in oral squamous cell carcinoma associated with betel quid chewing and cigarette smoking in Taiwanese. *Carcinogenesis* 22: 1497-1503.
- [68] Kouidou, S., Agidou, T., Kyrkou, A., Andreou, A., Katopodi, T., Georgiou, E., Krikelis, D., Dimitriadou, A., Spanos, P., Tsilikas, C., Destouni, H., and Tzimagiorgis G. 2005. Non-CpG cytosine methylation of p53 exon 5 in non-small cell lung carcinoma. *Lung Cancer*; 50: 299-307.
- [69] Rajesh, M., Wang, G., Jones, R., and Tretyakova, N. 2005. Stable isotope labeling-mass spectrometry analysis of methyl- and pyridyloxobutyl-guanine adducts of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in p53-derived DNA sequences. *Biochemistry* 44: 2197-2207.
- [70] S.Kouidou, A.Malousi, N.Maglaveras. Methylation and repeats in silent and nonsense mutations of p53, *Mutat. Res.* (2006).
- [71] S.Kouidou, A.Malousi, N.Maglaveras. Methylation and repeats in silent and nonsense mutations of p53, *Mutat. Res.* (2006).
- [72] Watson, R.E., Curtin, G.M., Doolittle, D.J., and Goodman, J.I. 2003. Progressive alterations in global and GC-rich DNA methylation during tumorigenesis. *Toxicol. Sci.* 75: 289-299.
- [73] Wostemeyer, J., and Kreibich, A. 2002. Repetitive DNA elements in fungi (Mycota): impact on genomic architecture and evolution. *Curr. Genet.* 41: 189-198.
- [74] Churchill, M.E., Jones, D.N., Glaser, T., Hefner, H., Searles, M.A. & Travers, A.A. (1995) HMG-D is an architecture-specific protein that preferentially binds to DNA containing the dinucleotide TG. *EMBO J.* 14, 1264±1275.
- [75] Calladine, C.R. & Drew, H.R. (1992) *Understanding DNA, the Molecule and How It Works.* Academic Press, London.
- [76] Structural bistability of repetitive DNA elements featuring CA/TG dinucleotide steps and mode of evolution of satellite DNA

[77] Laundon, C.H. & Griffith, J.D. (1988) Curved helix segments can uniquely orient the topology of supertwisted DNA. *Cell* 52, 545±549.

[78] Kato, M. (1993) Polypyrimidine/polypurine sequence in plasmid DNA enhances formation of dimer molecules in *Escherichia coli*. *Mol. Biol. Rep.* 18, 183±187.

[79] Rooney, S.M. & Moore, P.D. (1995) Antiparallel, intramolecular triplex DNA stimulates homologous recombination in human cells. *Proc. Natl Acad. Sci. USA* 92, 2141±2144.

[80] Wells, R.D. and Warren, S.T. (eds) (1998) *Genetic Instabilities and Hereditary Neurological Diseases*. Academic Press, San Diego.

[81] Sinden, R.R., Potaman, V.N., Oussatcheva, E.A., Pearson, C.E., Lyubchenko, Y.L. and Shlyakhtenko, L.S. (2002) Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J. Biosci.*, 27, 53–65.

[82] Bowater, R.P. and Wells, R.D. (2001) The intrinsically unstable life of DNA triplet repeats associated with human hereditary disorders. *Prog. Nucleic Acid Res. Mol. Biol.*, 66, 159–202.

[83] Napierala, M., Dere, R., Vetcher, A. and Wells, R.D. (2004) Structure-dependent recombination hot spot activity of GAA.TTC sequences from intron 1 of the Friedreich's ataxia gene. *J. Biol. Chem.*, 279, 6444–6454.

[84] Potaman, V.N., Bissler, J.J., Hashem, V.I., Oussatcheva, E.A., Lu, L., Shlyakhtenko, L.S., Lyubchenko, Y.L., Matsuura, T., Ashizawa, T., Leffak, M. et al. (2003) Unpaired structures in SCA10 (ATTCT)<sub>n</sub>(AGAAT)<sub>n</sub> repeats. *J. Mol. Biol.*, 326, 1095–1111.

[85] Siyanova, E.I. and Mirkin, S.M. (2001) Expansion of trinucleotide repeats. *Mol. Biol.*, 35, 208–223.

[86] Dere, R., Napierala, M., Ranum, L.P. and Wells, R.D. (2004) Hairpin structure-forming propensity of the (CCTG.CAGG) tetranucleotide repeats contributes to the genetic instability associated with myotonic dystrophy type 2. *J. Biol. Chem.*, 279, 41715–41726.

[87] Sakamoto, N., Chastain, P.D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J.D. and Wells, R.D. (1999) Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. *Mol. Cell*, 3, 465–475.

[88] Darlow, J.M. and Leach, D.R.F. (1998) Secondary structures in d(CGG) and d(CCG): repeat tracts. *J. Mol. Biol.*, 275, 3–16.

[89] Lahue, R.S. and Slater, D.L. (2003) DNA repair and trinucleotide repeat instability. *Front. Biosci.*, 8, 653–665.

[90] Hashem, V.I., Rosche, W.A. and Sinden, R.R. (2004) Genetic recombination destabilizes (CTG)<sub>n</sub>.(CAG)<sub>n</sub> repeats in *E. coli*. *Mutat. Res.*, 554, 95–109.

[91] Pearson, C.E., Ewel, A., Acharya, S., Fishel, R.A. and Sinden, R.R. (1997) Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum. Mol. Genet.*, 6, 1117–1123.

[92] Jaworski, A., Rosche, W.A., Gellibolian, R., Kang, S., Shimizu, M., Bowater, R.P., Sinden, R.R. and Wells, R.D. (1995) Mismatch repair in *Escherichia coli* enhances instability of (CTG)<sub>n</sub> triplet repeats from human hereditary diseases. *Proc. Natl Acad. Sci. USA*, 92, 11019–11023.

[93] Faruqi, A.F., Datta, H.J., Carroll, D., Seidman, M.M. and Glazer, P.M. (2000) Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway. *Mol. Cell. Biol.*, 20, 990–1000.

[94] Shigemori, Y. and Oishi, M. (2004) Specific cleavage of DNA molecules at RecA-mediated triple-strand structure. *Nucleic Acids Res.*, 32, e4.

[95] Vasquez, K.M., Christensen, J., Li, L., Finch, R.A. and Glazer, P.M. (2002) Human XPA and RPA DNA repair proteins participate in specific recognition of triplex-induced helical distortions. *Proc. Natl Acad. Sci. USA*, 99, 5848–5853.

[96] Robert D. Wells, Ruhee Dere, Micheal L. Hebert, Marek Napierala and Leslie S. Son. *Advances in mechanisms of genetic instability related to hereditary neurological diseases. Nucleic Acids Research*, 2005, Vol. 33, No. 12: 3785–3798.

[97] R.R. Sinden. *DNA Structure and Function*, Academic Press, San Diego, 1994.

[98] E.S. Lander, L.M. Linton, B. Birren, C. Nusbaum, M.C. Zody, J. Baldwin, K. Devon, K. Dewar, M. Doyle, W. FitzHugh, et al., Initial sequencing and analysis of the human genome, *Nature* 409 (2001) 860–921.

- [99] Guliang Wang, Karen M. Vasquez  
Non-B DNA structure-induced genetic instability. *Mutation Research* (2006)
- [100] D.K. Nag, T.D. Petes, Seven-base-pair inverted repeats in DNA form stable hairpins in vivo in *Saccharomyces cerevisiae*, *Genetics* 129 (1991) 669–673.
- [101] G.F. Richard, B. Dujon, J.E. Haber, Double-strand break repair can lead to high frequencies of deletions within shortCAG/CTG trinucleotide repeats, *Mol. Gen. Genet.* 261 (1999) 871–882.
- [102] J.L. Marcadier, C.E. Pearson, Fidelity of primate cell repair of a double-strand break within a (CTG)<sub>n</sub>(CAG)<sub>n</sub> tract. Effect of slipped DNA structures, *J. Biol. Chem.* 278 (2003) 33848–33856.
- [103] M.L. Hebert, L.A. Spitz, R.D. Wells, DNA double-strand breaks induce deletion of CTG·CAG repeats in an orientationdependent manner in *Escherichia coli*, *J. Mol. Biol.* 336 (2004) 655–672.
- [104] B. Malfoy, N. Rousseau, N. Vogt, E. Viegas-Pequignot, B. Dutrillaux, M. Leng, Nucleotide sequence of an heterochromatic segment recognized by the antibodies to Z-DNA in fixed metaphase chromosomes, *Nucl. Acids Res.* 14 (1986) 3197–3214.
- [105] B.H. Johnston, Generation and detection of Z-DNA, *Meth. Enzymol.* 211 (1992) 127–158.
- [106] H. Hamada, T. Kakunaga, Potential Z-DNA forming sequences are highly dispersed in the human genome, *Nature* 298 (1982) 396–398.
- [107] G.P. Schroth, P.J. Chou, P.S. Ho, Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes, *J. Biol. Chem.* 267 (1992) 11846–11855.
- [108] J. Klysik, S.M. Stirdivant, R.D. Wells, Left-handed DNA. Cloning, characterization, and instability of inserts containing different lengths of (dC–dG)<sub>n</sub> in *Escherichia coli*, *J. Biol. Chem.* 257 (1982) 10152–10158.
- [109] J. Klysik, S.M. Stirdivant, J.E. Larson, P.A. Hart, R.D. Wells,

Left-handed DNA in restriction fragments and a recombinant plasmid, *Nature* 290 (1981) 672–677.

[110] A.M. Freund, M. Bichara, R.P. Fuchs, Z-DNA-forming sequences are spontaneous deletion hot spots, *Proc. Natl. Acad. Sci. U.S.A.* 86 (1989) 7465–7469.

[111] B.Y. Ngan, J. Nourse, M.L. Cleary, Detection of chromosomal translocation t(14; 18) within the minor cluster region of bcl-2 by polymerase chain reaction and direct genomic sequencing of the enzymatically amplified DNA in follicular lymphomas, *Blood* 73 (1989) 1759–1762.

[112] J.R. Spitzner, I.K. Chung, M.T. Muller, Eukaryotic topoisomerase II preferentially cleaves alternating purine-pyrimidine repeats, *Nucl. Acids Res.* 18 (1990) 1–11.

[113] G.P. Schroth, P.S. Ho, Occurrence of potential cruciform and H-DNA forming sequences in genomic DNA, *Nucl. Acids Res.* 23 (1995) 1977–1983.

[114] B.S. Rao, Pausing of simian virus 40 DNA replication fork movement in vivo by (dG–dA)*n*·(dT–dC)*n* tracts, *Gene* 140 (1994) 233–237.

[115] M.M.Krasilnikova, S.M.Mirkin, Replication stalling at Friedreich's ataxia (GAA)*n* repeats in vivo, *Mol. Cell. Biol.* 24 (2004) 2286–2295.

[116] A Chompret<sup>1,3</sup>, L Brugières<sup>1</sup>, M Ronsin<sup>2</sup>, M Gardes<sup>2</sup>, F Dessarps-Freichy<sup>2</sup>, A Abel<sup>1,3</sup>, D Hua<sup>3</sup>, L Ligot<sup>3</sup>, M-G Dondon<sup>3</sup>, B Bressac-de Paillerets<sup>4</sup>, T Frébourg<sup>5</sup>, J Lemerle<sup>1</sup>, C Bonaïti-Pellié<sup>3</sup> and J Feunteun. P53 germline mutations in childhood cancers and cancer risk for carrier individuals. *British Journal of Cancer* (2000) 82(12), 1932–1937

[117] Sally G. Warneford,<sup>2</sup> Leah J. Witton, Megan L Townsend, Peter B. Rowe, Roger R. Reddel, Luciano DaIia-Pozza, and Geoff Symonds. Germ-Line Splicing Mutation of the p53 Gene in a Cancer-prone Family. *Cell Growth & Differentiation* Vol. 3, 839-846, 1992.

[118] JM Varley, G McGown, M Thorncroft, M Santibanez-Koref, A Kelsey, KJ Tricker, D Gareth, R Evans, JM Birch. Germline mutations of tp53 in Li Fraumeni families: an extended study of 39 families. *Cancer Research* 57, 3245-3252, 1997.

- [119] Li, F. P., and Fraumeni, J. F. Soft-tissue sarcomas, breast cancer, and other neoplasms, a familial syndrome? *Ann. Intern. Med.*, *71*: 747–752, 1969.
- [120] Li, F. P., Fraumeni, J. F., Mulvihill, J. J., Blattner, W. A., Dreyfus, M. G., Tucker, M. A., and Miller, R. W. A cancer family syndrome in twenty-four kindreds. *Cancer Res.*, *48*: 5358–5362, 1988.
- [121] Garber, J. E., Burke, E. M., Lavalley, B. L., Billett, A. L., Sallan, S. E., Scott, R. M., Kupsky, W., and Li, F. P. Choroid plexus tumors in the breast cancer sarcoma syndrome. *Cancer (Phila.)*, *66*: 2658–2660, 1990.
- [122] Garber, J. E., Goldstein, A. M., Kantor, A. F., Dreyfus, M. G., Fraumeni, J. F., and Li, F. P. Follow-up study of twenty-four families with Li-Fraumeni syndrome. *Cancer Res.*, *51*: 6094–6097, 1991.
- [123] Hisada, M., Garber, J. E., Fung, C. Y., Fraumeni, J. F., and Li, F. P. Multiple primary cancers in families with Li-Fraumeni Syndrome. *J. Natl. Cancer Inst. (Bethesda)*, *90*: 606–611, 1998.
- [124] Malkin, D., Li, F. P., Strong, L. C., Fraumeni, J. F., Nelson, C. E., Kim, D. H., Kassel, J., Gryka, M. A., Bischoff, F. Z., Tainsky, M. A., and Friend, S. H. Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science (Washington DC)*, *250*: 1233–1238, 1990.
- [125] Frebourg, T., Barbier, N., Yan, Y-X., Garber, J. E., Dreyfus, M., Fraumeni, J., Li, F. P., and Friend, S. H. Germ-line p53 mutations in 15 families with Li-Fraumeni syndrome. *Am. J. Hum. Genet.*, *56*: 608–615, 1995.
- [126] Diller, L., Sexsmith, E., Gottleib, A., Li, F. P., and Malkin, D. Germline p53 mutations are frequently detected in young children with rhabdomyosarcoma. *J. Clin. Investig.*, *95*: 1606–1611, 1995.
- [127] Claire Morgan, Paul D. Lewis. iMARS—Mutation analysis reporting software: An analysis of spontaneous *cII* mutation spectra. *Mutation Research* 603 (2006) 15–26
- [128] D.N. Cooper, M. Krawczak, The mutational spectrum of single base-pair substitutions causing human genetic disease: patterns and predictions, *Hum. Genet.* 85 (1990) 55–74.
- [129] R.E. Tarone, Testing for non-randomness of events in sparse data situations, *Ann. Hum. Genet.* 53 (1989) 381–387.



[130] W.T. Adams, T.R. Skopek, Statistical test for the comparison of samples from mutational spectra, *J. Mol. Biol.* 194 (1987) 391–396.

[131] I.B. Rogozin, F.A. Kondrashov, G.V. Glazko, Use of mutation spectra analysis software, *Hum. Mutat.* 17 (2001) 83–201.

[132] Cancer Statistics Review 1975-2002, National Cancer Institute, USA

## ΠΑΡΑΡΤΗΜΑ

### Ο Γενετικός Κώδικας

Τα αμινοξέα και τα σύμβολά τους			Κωδικόνια						
ασπαρτικό οξύ	Asp	D	GAC	GAU					
γλουταμικό οξύ	Glu	E	GAA	GAG					
αργινίνη	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
λυσίνη	Lys	K	AAA	AAG					
ιστιδίνη	His	H	CAC	CAU					
ασπαραγίνη	Asn	N	AAC	AAU					
γλουταμίνη	Gln	Q	CAA	CAG					
σερίνη	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
θρεονίνη	Thr	T	ACA	ACC	ACG	ACU			
τυροσίνη	Tyr	Y	UAC	UAU					
αλανίνη	Ala	A	GCA	GCC	GCG	GCU			
γλυκίνη	Gly	G	GGA	GGC	GGG	GGU			
βαλίνη	Val	V	GUA	GUC	GUG	GUU			
λευκίνη	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
ισολευκίνη	Ile	I	AUA	AUC	AUU				
προλίνη	Pro	P	CCA	CCC	CCG	CCU			
φαινυλαλανίνη	Phe	F	UUC	UUU					
μεθειονίνη	Met	M	AUG						
τρυπτοφάνη	Trp	W	UGG						
κυστεΐνη	Cys	C	UGC	UGU					
			UAA	UAG	UGA				

Κωδικόνια ΤΕΡΜΑΤΙΣΜΟΥ

ΚΛΕΙΔΙ:  αρνητικά φορτισμένα πολικά αμινοξέα  θετικά φορτισμένα πολικά αμινοξέα  
 μη φορτισμένα πολικά αμινοξέα  μη πολικά αμινοξέα

## Ξενόγλωσση Ορολογία

**Somatic** – σωματικά κύτταρα

**Germline** – γενετικά κύτταρα

**Germline per family** – γενετικές μεταλλάξεις με μία καταγραφή ανά οικογένεια στον πίνακα δεδομένων

**Silent** – σιωπηρές μεταλλάξεις, δηλαδή δίχως να προκαλούν αλλαγή του αμινοξέως

**Nonsense** – μεταλλάξεις που οδηγούν στον τερματισμό της πρωτεϊνοσύνθεσης (τερματισμού)

**Missense** – μεταλλάξεις που προκαλούν αλλαγή του αμινοξέως στην πρωτεΐνη (παρανοηματικές)

**FS (Frameshift)** – προσθήκη ή αφαίρεση βάσεων που έχει αποτέλεσμα την μετακίνηση του μεταγραφικού πλαισίου ανάγνωσης

**NA** – μεταλλάξεις για τις οποίες δεν έχουμε στοιχεία για το χαρακτηριστικό που μελετάμε

**Splice** – μεταλλάξεις που συμβαίνουν μέσα σε ενδόνια και επηρεάζουν το splicing (ματίσματος)

**Intronic** – μεταλλάξεις που συμβαίνουν μέσα σε ενδόνια και δεν επηρεάζουν το μάτισμα και αφήνουν αμετάβλητη την τελική πρωτεΐνη

**Other** – μεταλλάξεις που δεν μπορούν να προσδιοριστούν από άλλη κατηγορία

**Tandem mutations** – επαναλαμβανόμενες μεταλλάξεις

**Hotspots** – είναι οι γενετικές θέσεις του γονιδίου που παρουσιάζουν αυξημένη μεταλλαξιμότητα

**Maternal** – μητρική κληρονομία γονιδίου

**Paternal** – πατρική κληρονομία γονιδίου

**LFS** – Li Fraumeni σύνδρομο

**LFL** – Li Fraumeni Like σύνδρομο

**NO FH** – No Family History (Δεν υπάρχει οικογενειακό ιστορικό)

**FH** – Family History (Υπάρχει οικογενειακό ιστορικό)