

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ**

**ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ**

**ΤΜΗΜΑ ΙΑΤΡΙΚΗΣ**



**ΕΡΓΑΣΤΗΡΙΟ ΚΥΤΤΑΡΟΓΕΝΕΤΙΚΗΣ ΚΑΙ ΜΟΡΙΑΚΗΣ ΓΕΝΕΤΙΚΗΣ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ**

**«ΓΕΝΕΤΙΚΗ ΤΟΥ ΑΝΘΡΩΠΟΥ»**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**«ΚΛΙΝΙΚΕΣ ΕΦΑΡΜΟΓΕΣ ΤΗΣ ΑΛΛΗΛΟΥΧΗΣΗΣ ΝΕΑΣ ΓΕΝΙΑΣ»**

**ΜΑΡΙΑ ΣΤΑΜΑΤΟΠΟΥΛΟΥ**

**ΙΑΤΡΟΣ**

**ΛΑΡΙΣΑ**

**ΙΟΥΝΙΟΣ 2014**

ΔΙΕΥΘΥΝΤΡΙΑ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ  
*ΚΑΘΗΓΗΤΡΙΑ Α. ΤΣΕΖΟΥ*

**ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ**

**Τσέζου Ασπασία**, Καθηγήτρια Ιατρικής Γενετικής Ιατρικού Τμήματος Πανεπιστημίου Θεσσαλίας, Διευθύντρια του Εργαστηρίου Κυτταρογενετικής και Μοριακής Γενετικής του ΠΓΝΛ , Διευθύντρια του Προγράμματος Μεταπτυχιακών Σπουδών «Γενετική του Ανθρώπου» του Ιατρικού Τμήματος του Πανεπιστημίου Θεσσαλίας (Επιβλέπουσα)

**Τζέτη Μαρία**, Επίκουρη Καθηγήτρια Εργαστηρίου Ιατρικής Γενετικής της Ιατρικής Σχολής Πανεπιστημίου Αθηνών (Συνεπιβλέπουσα)

**Κυριάκου Δέσποινα**, Αναπληρώτρια Καθηγήτρια Αιματολογίας ΠΓΝΛ, Διευθύντρια Αιμοδοσίας-Αιματολογικού Τμήματος Π.Γ.Ν. Λάρισας (Συνεπιβλέπουσα)

Ευχαριστώ ιδιαίτερος την Επικ.Καθηγήτρια ΕΚΠΑ κ.Τζέτη Μαρία-συνεπιβλέπουσα, για την καθοδήγηση της στην εκπόνηση της παρούσης μεταπτυχιακής εργασίας.



**ΚΛΙΝΙΚΕΣ ΕΦΑΡΜΟΓΕΣ  
ΤΗΣ  
ΑΛΛΗΛΟΥΧΗΣΗΣ ΝΕΑΣ ΓΕΝΙΑΣ**

# Πίνακας Περιεχομένων

|                                                                                                              |           |
|--------------------------------------------------------------------------------------------------------------|-----------|
| Πίνακας Περιεχομένων .....                                                                                   | 5         |
| Περίληψη .....                                                                                               | 7         |
| Abstract .....                                                                                               | 8         |
| Εισαγωγή .....                                                                                               | 9         |
| <b>2. Μεθοδολογία .....</b>                                                                                  | <b>10</b> |
| 2.1 Γενικά.....                                                                                              | 10        |
| <b>3. Κύριο μέρος βιβλιογραφικής επισκόπησης .....</b>                                                       | <b>11</b> |
| 3.1 Ιστορική επισκόπηση της αλληλούχησης.....                                                                | 11        |
| 3.1.1 Η μέθοδος Sanger .....                                                                                 | 12        |
| 3.1.2. Η μέθοδος της χημικής αποδόμησης των Maxam-Gilbert .....                                              | 13        |
| 3.2 Η αλληλούχηση του ανθρώπινου γονιδιώματος .....                                                          | 16        |
| 3.3 Μεταλλάξεις .....                                                                                        | 17        |
| 3.3.1 Γενικά στοιχεία.....                                                                                   | 17        |
| 3.4 Αλληλούχηση Επόμενης Γενιάς (Next Generation Sequencing).....                                            | 18        |
| 3.4.1 Γενικά στοιχεία.....                                                                                   | 18        |
| 3.4.2 Genome Sequencer FLX System (454).....                                                                 | 20        |
| 3.4.3 The Illumina (Solexa) Genome Analyzer .....                                                            | 22        |
| 3.4.4 The Applied Biosystem SOLiD system .....                                                               | 25        |
| 3.5 Εφαρμογές.....                                                                                           | 27        |
| 3.5.1 Γενικά στοιχεία.....                                                                                   | 27        |
| 3.5.2 Αλληλούχηση επόμενης γενιάς και συγγενείς διαταραχές γλυκοζυλίωσης ....                                | 28        |
| 3.5.3 Αλληλούχηση επόμενης γενιάς και καρκίνος .....                                                         | 29        |
| 3.5.4 Περιγραφή των κλινικών εξετάσεων για την αντιμετώπισης του καρκίνου με<br>χρήση της τεχνικής NGS ..... | 33        |
| 3.5.4 Αλληλούχηση επόμενης γενιάς και ασθένειες που σχετίζονται με τη<br>μιτοχονδριακή λειτουργία.....       | 35        |
| 3.5.5 Αλληλούχηση επόμενης γενιάς και καρδιαγγειακές παθήσεις.....                                           | 36        |

|                                                                           |           |
|---------------------------------------------------------------------------|-----------|
| 3.5.6 Αλληλούχηση επόμενης γενιάς και αμφιβληστροειδοπάθειες.....         | 37        |
| 3.5.7 Αλληλούχηση επόμενης γενιάς και αιματολογικές διαταραχές.....       | 38        |
| 3.5.8 Αλληλούχηση επόμενης γενιάς και πρωτοπαθής δυσκινησία των κροσσών . | 39        |
| 3.5.9 Αλληλούχηση επόμενης γενιάς και επιληψία .....                      | 40        |
| 3.5.10 Αλληλούχηση επόμενης γενιάς και διαταραχές της ακοής.....          | 41        |
| 3.6 Αλληλούχηση επόμενης γενιάς και επιγενετική.....                      | 43        |
| 3.6.1.Γενικά στοιχεία.....                                                | 43        |
| 3.7 Πρωτογενής ανάλυση δεδομένων αλληλούχησης επόμενης γενιάς .....       | 45        |
| 3.7.1.Γενικά στοιχεία.....                                                | 45        |
| <b>4. Συζήτηση – Συμπεράσματα .....</b>                                   | <b>47</b> |
| <b>5. Βιβλιογραφία .....</b>                                              | <b>50</b> |

# Περίληψη

Η αλληλούχιση του DNA και κατ' επέκταση η αποκωδικοποίηση του ανθρώπινου γονιδιώματος υπήρξαν ορόσημα στην Ιατρική, διότι έγινε γνωστή η αλληλουχία όλων των γονιδίων του ανθρώπινου οργανισμού. Τα τελευταία χρόνια οι τεχνολογίες υψηλής απόδοσης (High throughput technologies) και η αλληλούχιση επόμενης γενιάς (Next Generation Sequencing – NGS) μπορούν να αποτελέσουν πολύ σημαντικά εργαλεία για την κλινική διάγνωση πολλών ασθενειών και να διευκολύνουν την ανάπτυξη και την εισαγωγή νέων και εξατομικευμένων θεραπειών. Μέσω της αλληλούχισης επόμενης γενιάς γίνεται δυνατό με μικρό κόστος και σε σύντομο χρονικό διάστημα να αναλυθεί η αλληλουχία πάνελ πολυάριθμων γονιδίων που σχετίζονται με συγκεκριμένες νόσους, ενώ εφαρμόζεται στην επανααλληλούχιση γονιδιωμάτων, στο μεταγραφικό προφίλ, στις αλληλεπιδράσεις του DNA με πρωτεΐνες και στο χαρακτηρισμό του επιγενώματος. Είναι δεδομένο πως στο μέλλον η αλληλούχιση νέας γενιάς θα χρησιμοποιείται ευρέως στις κλινικές εφαρμογές και θα δίνει πολύ ξεχωριστές πληροφορίες για τον κάθε ασθενή αλλά και σε ομάδες αυτών.

Λέξεις κλειδιά: **Αλληλούχιση, αλληλούχιση νέας γενιάς, κλινικές εφαρμογές αλληλούχισης, αλληλούχιση στη ιατρική,**

# Abstract

Human genome sequencing was a milestone of medicine research as the whole genome was sequenced and became available. Lately, high throughput technologies and Next Generation Sequencing – NGS, have become invaluable tools in clinical applications because of their utility in diagnosis and personalized treatments. This technique applies to genome sequencing, genome resequencing, transcriptome profiling (RNA-Seq), DNA-protein interactions (ChIP-sequencing), and epigenome characterization. The recent development of high-throughput next-generation sequencing (NGS) technology has transformed the way DNA-based molecular diagnostic testing is performed in clinical laboratories, as NGS allows parallel sequencing analyses of multiple genes effectively at any desirable depth of coverage and gives unprecedented information for gene-related diseases.

Keywords: **Sequencing, Next Generation Sequencing, sequencing and clinical applications, Next Generation Sequencing technologies in medical genetics, Next Generation Sequencing and clinical diagnostics**



# Εισαγωγή

Η αλληλούχιση επόμενης γενιάς (Next Generation Sequencing – NGS), οι πρωτοποριακές τεχνολογίες υψηλής απόδοσης (High throughput technologies) σε συνδυασμό με τις τεχνολογίες συστοιχιών (array-based methods), έχουν τη δυνατότητα να δώσουν ολοκληρωμένες πληροφορίες για το ανθρώπινο γονιδίωμα.

Τα τελευταία χρόνια έχει αρχίσει η ερευνητική εφαρμογή της αλληλούχισης επόμενης γενιάς σε κλινικά δείγματα με απώτερο στόχο τη γονιδιωματική διαγνωστική, αλλά και τη θεραπευτική προσέγγιση σε νόσους όπως είναι ο καρκίνος, ο σακχαρώδης διαβήτης, οι νευροαναπτυξιακές νόσοι όπως η σχιζοφρένεια και ο αυτισμός, καθώς και σε άλλες πολυπαραγοντικές νόσους, καθώς είναι δεδομένο πως οι αλλαγές στη δομή και τη λειτουργία του γονιδιώματος οδηγούν στην εμφάνιση του καρκίνου και άλλων πολύ σοβαρών παθήσεων. Η αλληλούχιση επόμενης γενιάς στο ανθρώπινο γονιδίωμα παρέχει τη μοναδική δυνατότητα αναγνώρισης όλων των ειδών μεταλλάξεων, σημειακών μεταλλάξεων και πολυμορφισμών (SNPs), ή και μεγαλύτερων διαταραχών όπως οι αλλαγές στον αριθμό των αντιγράφων και τις χρωμοσωμικές αναδιατάξεις. Λόγω της ετερογένειας των μεταλλάξεων από ασθενή σε ασθενή με την ίδια νόσο και τα ίδια κλινικο-παθολογοανατομικά χαρακτηριστικά, η σημαντικότητα της αλληλούχισης επόμενης γενιάς είναι αδιαμφισβήτητη.

Στόχος της παρούσας πτυχιακής μελέτης είναι η περιγραφή της καινοτόμου τεχνολογίας του Next Generation Sequencing ανάλογα με την εμπορική πλατφόρμα καθώς και η παρουσίαση των κλινικών εφαρμογών της, ιδιαίτερα τα τελευταία χρόνια. Αρχικά παρουσιάζεται μια ιστορική διαδρομή της αλληλούχισης και των πρώτων κλασσικών μεθόδων και στη συνέχεια περιγράφεται η αλληλούχιση του ανθρώπινου γονιδιώματος το 2001 που αποτέλεσε σημαντικό ορόσημο στην Ιατρική και τη Βιολογία γενικότερα. Στη συνέχεια γίνεται περιγραφή των μεταλλάξεων και των διαταραχών που παρουσιάζονται στο ανθρώπινο γονιδίωμα ώστε να μπορούν να γίνουν κατανοητές οι τεχνολογίες αλληλούχισης. Περιγράφονται οι πιο σημαντικές εμπορικές πλατφόρμες και οι κλινικές εφαρμογές τους στον καρκίνο, στις συγγενείς διαταραχές γλυκοζυλίωσης, σε ασθένειες που σχετίζονται με τη μιτοχονδριακή λειτουργία, στις αμφιβληστροειδοπάθειες και τις καρδιαγγειακές παθήσεις. Συνοψίζοντας δίνονται τα συμπεράσματα από την ανάλυση της βιβλιογραφίας και δίνονται συνοπτικά με

παραδείγματα οι εφαρμογές του Next Generation Sequencing στην Ιατρική.

## 2. Μεθοδολογία

### 2.1 Γενικά

Οι λέξεις «κλειδιά» που χρησιμοποιήθηκαν είτε ξεχωριστά η κάθε μία, είτε σε συνδυασμό μεταξύ τους για να βρεθούν τα άρθρα, είναι οι εξής:

- στην ελληνική: **Αλληλούχιση, Αλληλούχιση νέας γενιάς, κλινικές εφαρμογές αλληλούχισης, αλληλούχιση στη ιατρική**
- στην αγγλική: **Sequencing, Next generation sequencing, sequencing and clinical applications, Next Generation Sequencing technologies in medical genetics, Next Generation Sequencing and clinical diagnostics**

Οι βάσεις δεδομένων που χρησιμοποιήθηκαν είναι οι PubMed, Scopus και Google Scholar. Έγινε χρήση όλων των διαθέσιμων φίλτρων και εργαλείων καθώς έγινε και αναζήτηση μέσω των αναφορών σε άλλα άρθρα. Αναζητήθηκαν είτε ερευνητικά και πειραματικά άρθρα (Research papers), είτε άρθρα ανασκόπησης (Reviews). Επίσης, έγινε αναζήτηση ως προς τα ονόματα συγγραφέων που ασχολούνται με το συγκεκριμένο επιστημονικό αντικείμενο.

Με την εξειδικευμένη αναζήτηση μέσω των παραπάνω λέξεων σε συνδυασμό με τα φίλτρα, πραγματοποιήθηκε η βιβλιογραφική ανασκόπηση που αφορά στην αλληλούχιση επόμενης γενιάς (Next generation sequencing) και τις κλινικές εφαρμογές. Τα άρθρα που βρέθηκαν αναλύονται αναλυτικά στο κύριο μέρος της βιβλιογραφικής επισκόπησης.

## 3. Κύριο μέρος βιβλιογραφικής επισκόπησης

### 3.1 Ιστορική επισκόπηση της αλληλούχισης

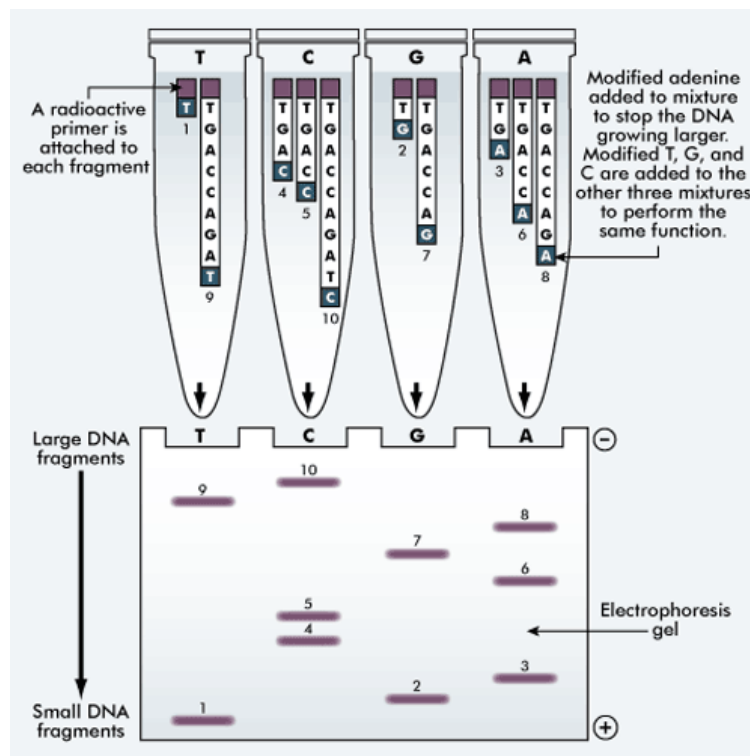
Η πρώτη αλληλούχιση DNA (1968) πραγματοποιήθηκε 15 χρόνια αργότερα από την ανακάλυψη της διπλής έλικας το 1953, (Hutchison, 2007). Παρόλα αυτά η χημική μέθοδος των Maxam και Gilbert και η μέθοδος του Sanger ξεκίνησαν και αναπτύχθηκαν στα μέσα της δεκαετίας του 1970. Σημαντικές πληροφορίες για τη γενετική οργάνωση παρουσιάσθηκαν από τους Nicklen και Coulson όταν η πρώτη ολόκληρη ακολουθία αλληλουχήθηκε από το φάγο φΧ174. Όταν οι μέθοδοι αλληλούχισης βελτιώθηκαν, μεγαλύτερα μόρια μεγέθους άνω των 200kb (ανθρώπινος κυτταρομεγαλοϊός, human cytomegalovirus-CMV) αλληλουχήθηκαν με αποτέλεσμα να δημιουργηθεί και η ανάγκη για βιοπληροφορική ανάλυση των δεδομένων. Οι προσπάθειες για αλληλούχιση υποστηρίχθηκαν σημαντικά με το πρόγραμμα αλληλούχισης του ανθρώπινου γονιδιώματος (US Human Genome Project) από το 1992 (Hutchison, 2007). Παράλληλα με αυτή την προσπάθεια πραγματοποιήθηκε και η αλληλούχιση του πρώτου βακτηριακού γονιδιώματος το 1995, ενώ και το γονιδίωμα άλλων αρχαιοβακτηριών και ευκαρυωτικών οργανισμών αλληλουχήθηκε στη συνέχεια. Το αποτέλεσμα του ανταγωνισμού μεταξύ της εταιρείας Celera genomics και του δημοσίου προγράμματος του Αμερικανικού Ινστιτούτου Υγείας (US NIH Human Genome Project), ήταν η δημοσίευση του ανθρώπινου γονιδιώματος το 2001. Οι καινούργιες μέθοδοι μαζικής αλληλούχισης αύξησαν σημαντικά στη συνέχεια τις δυνατότητες, αλλά χρειάζονται κι άλλες καινοτομίες προκειμένου να επιτευχθεί το “γονιδίωμα των 1000 δολαρίων” (thousand dollar genome), το οποίο πολλοί θεωρούν ως προαπαιτούμενο για την εξατομικευμένη γενετική ιατρική. Βέβαια, αυτές οι πρόοδοι θα επιτρέψουν επίσης νέες προσεγγίσεις για μια ποικιλία προβλημάτων στη βιολογία, την εξέλιξη και το περιβάλλον.

Το 1975, ο Sanger εισήγαγε τη μέθοδο αλληλούχισης DNA “plus and minus”, (Sanger and Coulson, 1975). Αυτή η τεχνική ήταν σημαντική για τη μετάβαση στις μοντέρνες τεχνικές που καθιερώθηκαν για περισσότερα από 30 χρόνια. Το κλειδί για αυτή την εξέλιξη ήταν η χρησιμοποίηση gels πολυακρυλαμίδης με τα οποία ήταν δυνατό να διαχωριστούν τα προϊόντα σύνθεσης DNA από πολυμεράση η οποία αύξανε

την νουκλεοτιδική αλυσίδα (Hutchison, 2007).

### **3.1.1 Η μέθοδος Sanger**

Στη μέθοδο του Sanger, χρησιμοποιείται μια DNA εξαρτώμενη πολυμεράση για να δημιουργήσει ένα συμπληρωματικό μονόκλωνο τμήμα DNA (αλληλούχιση μέσω σύνθεσης-Sequencing by synthesis, SBS) (Sanger et al., 1988). Ξεκινώντας από το 3' άκρο, συντίθεται μια νέα συμπληρωματική μονόκλωνη αλυσίδα DNA. Δεοξυνουκλεοτίδια προστίθενται στην αυξανόμενη αλυσίδα και είναι συμπληρωματικά των νουκλεοτιδίων της μήτρας του DNA (Εικόνα 1). Η δημιουργία ενός φωσφοδιεστερικού δεσμού μεταξύ του 3' υδροξυλίου και της 5' φωσφορικής ομάδας με το εισερχόμενο δεοξυνουκλεοτίδιο μπορεί να επιμηκύνει την αλυσίδα του DNA και να απελευθερώνει ένα πυροφωσφορικό (PPi) (Shendure et al., 2011).



**Εικόνα 1: Διαγραμματική απεικόνιση της μεθόδου Sanger**

([http://tle.westone.wa.gov.au/content/file/969144ed-0d3b-fa04-2e88-8b23de2a630c/1/human\\_bio\\_science\\_3b.zip/content/005\\_dna/page\\_15.htm](http://tle.westone.wa.gov.au/content/file/969144ed-0d3b-fa04-2e88-8b23de2a630c/1/human_bio_science_3b.zip/content/005_dna/page_15.htm))

Η διαφορά μεταξύ των 2',3' ddNTPs από τα κανονικά δεοξυ-τριφωσφορικά νουκλεοτίδια (dNTPs) είναι ότι από τα πρώτα λείπει ένα υδροξύλιο από την 3' θέση της

δεοξυριβόζης. Μπορούν δε να εισαχθούν με μια αντίδραση πολυμερισμού με μια DNA πολυμεράση, μέσω της 5' τριφωσφορικής τους ομάδας, σε μια αλυσίδα DNA. Παρόλα αυτά η έλλειψη του υδροξυλίου στην 3' δέση δεν επιτρέπει στη συνέχεια το σχηματισμό φωσφοροδιεστερικού δεσμού με το επόμενο δεοξυνουκλεοτίδιο. Έτσι δεν είναι δυνατό να συνεχιστεί παραπέρα ο πολυμερισμός της αλυσίδας του DNA. Όταν σε μια αντίδραση πολυμερισμού του DNA, μαζί με τα τέσσερα κανονικά νουκλεοτίδια (dNTPs) περιλαμβάνεται και μια μικρή ποσότητα ενός από τα 4 διδεόξυ-τριφωσφορικά νουκλεοτίδια (ddNTPs), υπάρχει ανταγωνισμός μεταξύ πολυμερισμού της αλυσίδας και του λιγότερο συχνού, αλλά εξειδικευμένου τερματισμού. Τα προϊόντα μιας τέτοιας αντίδρασης είναι μια σειρά ολιγονουκλεοτιδικών αλυσίδων, το μήκος των οποίων καθορίζεται από την απόσταση από την αρχή του εκκινήτη που χρησιμοποιείται για την έναρξη του πολυμερισμού και τη θέση του τερματισμού. Με την χρήση των τεσσάρων διαφορετικών ddNTPs σε 4 διαφορετικές ενζυμικές αντιδράσεις, επιτυγχάνεται η σύνδεση ολιγονουκλεοτιδίων που τερματίζουν με καθένα από τα 4 νουκλεοτίδια A, G, C, T. Σαν αλυσίδα μήτρα για τον πολυμερισμό χρησιμοποιούνται, είτε μονόκλωνα μόρια DNA είτε αποδιαταγμένες δίκλωνες αλυσίδες. Σαν εκκινήτες χρησιμοποιούνται συνθετικά ολιγονουκλεοτίδια που είναι συμπληρωματικά με μια συγκεκριμένη ακολουθία στην αλυσίδα μήτρα. Τα ένζυμα που χρησιμοποιούνται στον πολυμερισμό με την μέθοδο Sanger (Sanger et al., 1977), είναι το τμήμα Klenow της *E.coli* πολυμεράσης I (Sanger et al. 1977), η ανάστροφη μεταγραφάση (Mierendorf and Pfeffer, 1987) και οι T7 DNA πολυμεράσες που έχουν τροποποιηθεί έτσι ώστε έχει καταστραφεί η 3'->5' εξωνουκλεολυτική τους δράση. Τέτοιες πολυμεράσες είναι η Sequense και η Sequense version 2.0 (Tabor and Richardson, 1987). Χρησιμοποιείται επίσης η θερμοσταθερή DNA πολυμεράση που έχει απομονωθεί από τον *Thermus aquaticus* (*Taq* DNA πολυμεράση). Οι ιδιότητες και τα χαρακτηριστικά των ενζύμων αυτών διαφέρουν μεταξύ τους κατά τέτοιο τρόπο που επηρεάζουν την ποιότητα και το μήκος της ακολουθίας που συντίθεται.

### **3.1.2. Η μέθοδος της χημικής αποδόμησης των Maxam-Gilbert**

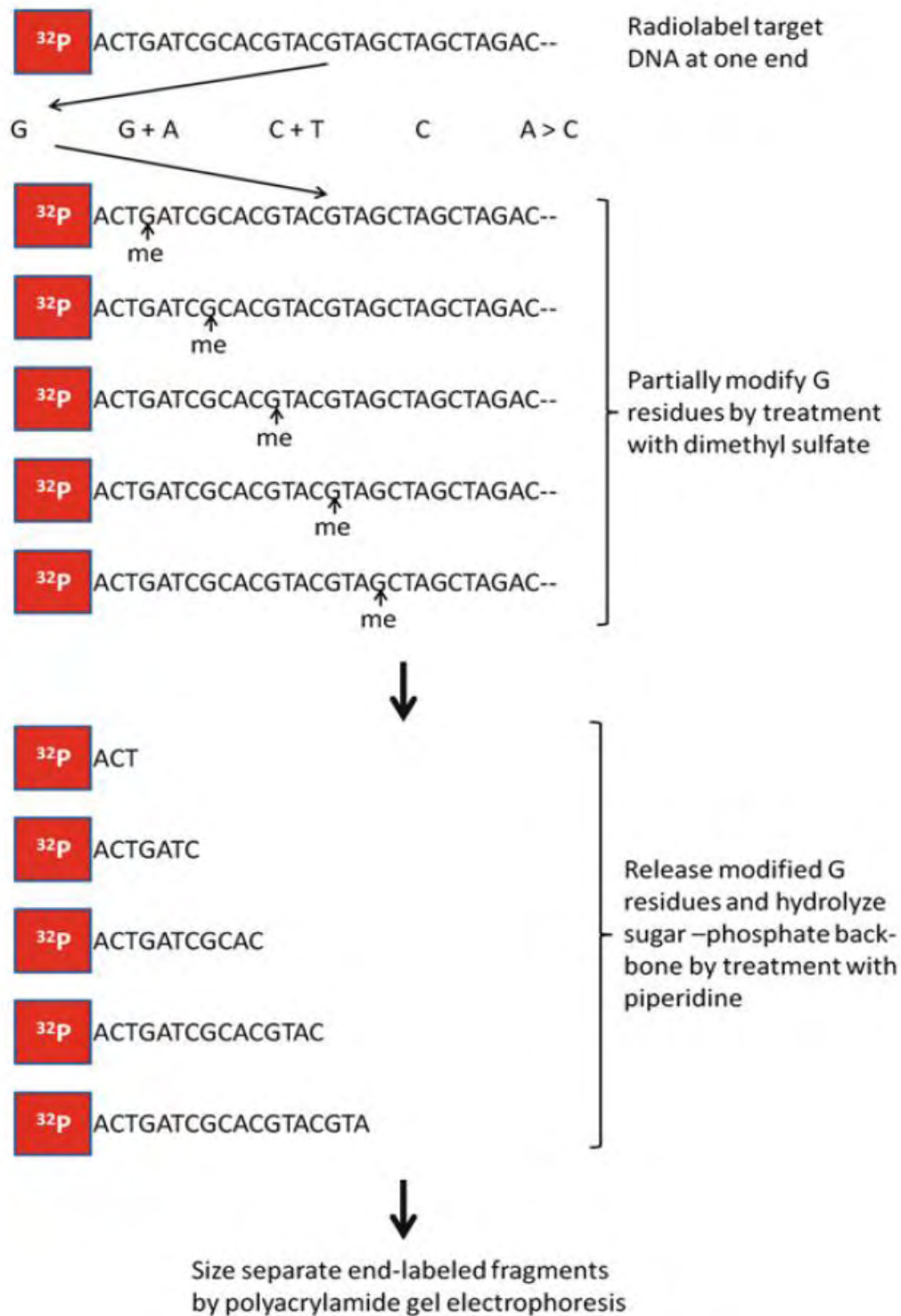
Αντίθετα με τη μέθοδο του τερματισμού της αντίδρασης στην οποία γίνεται σύνθεση της αλυσίδας του DNA, με τη μέθοδο Maxam-Gilbert λαμβάνει μέρος αποδόμηση της αρχικής ακολουθίας του DNA. Στη μέθοδο αυτή (Maxam and Gilbert,

1980) ένα τμήμα DNA ραδιοσημασμένο στο ένα άκρο του σπάει σε τέσσερις ξεχωριστές χημικές αντιδράσεις, καθεμία από τις οποίες είναι ειδική για μια συγκεκριμένη βάση ή τύπο βάσης. Έτσι δημιουργούνται τέσσερις πληθυσμοί ραδιοσημασμένων μορίων που ξεκινούν από ένα κοινό σημείο (το ραδιοσημασμένο άκρο), μέχρι τη θέση που έχει συμβεί το κόψιμο (Εικόνα 2). Κάθε πληθυσμός αποτελείται από ένα μίγμα μορίων, το μήκος των οποίων καθορίζεται από τη θέση των συγκεκριμένων βάσεων που συμβαίνει το κόψιμο στην αρχική αλυσίδα του DNA. Οι πληθυσμοί αυτοί αναλύονται σε αποδιατακτική ηλεκτροφόρηση σε πηκτή πολυακρυλαμίδης και τα ραδιοσημασμένα μόρια ανιχνεύονται με αυτοραδιογραφία.

Η μέθοδος Maxam-Gilbert είναι λιγότερο αποδοτική όσον αφορά το μέγεθος της ακολουθίας που μπορεί να αναλυθεί. Όταν όμως στη δεκαετία του '70 αναπτύχθηκαν οι δύο τεχνικές, το sequencing με χημική αποδόμηση ήταν πιο αποδοτικό και πιο εύκολο στην εφαρμογή του, αφού η τεχνική του Sanger απαιτούσε μονόκλωνη μήτρα, εξειδικευμένους εκκινητές και μεγάλης καθαρότητας ένζυμο ενώ αυτή των Maxam-Gilbert χρησιμοποιούσε χημικούς παράγοντες με εύκολη πρόσβαση στους ερευνητές.

Παρόλα αυτά με τη δημιουργία του βακτηριοφάγου M13 και των phagemids, καθώς και μηχανημάτων που κάνουν εύκολη και φθηνή την κατασκευή συνθετικών εκκινητών και διάφορες άλλες βελτιώσεις που έχουν γίνει στις αντιδράσεις του sequencing, η μέθοδος του Sanger χρησιμοποιείται σήμερα σε ευρεία κλίμακα.

Η μέθοδος όμως της χημικής αποδόμησης έχει ένα σημαντικό και ξεκάθαρο πλεονέκτημα σε σχέση με τη μέθοδο τερματισμού της αντίδρασης. Η ακολουθία προέρχεται από την αρχική DNA μήτρα και δεν είναι αντίγραφο αυτής, με ό,τι πιθανά λάθη δύναται να συνεπάγεται αυτό.



**Εικόνα 2:** Διαγραμματική απεικόνιση της μεθόδου Maxam-Gilbert (Wong, 2013)

## 3.2 Η αλληλούχιση του ανθρώπινου γονιδιώματος

Η αλληλούχιση του ανθρώπινου γονιδιώματος υπήρξε ένας από τους σημαντικότερους στόχους της ιατρικής επιστήμης. Επίσημες συζητήσεις για την επίτευξη του στόχου ξεκίνησαν από το 1985 όταν ο Robert Sinsheimer οργάνωσε ένα συνέδριο για την αλληλούχιση του ανθρώπινου γονιδιώματος στο Πανεπιστήμιο της Καλιφόρνια στο Santa-Cruz (Sinsheimer, 2006). Το ίδιο έτος οι Charles DeLisi και David A. Smith οργάνωσαν το πρώτο συνέδριο με την υποστήριξη του DOE (Department of Energy), για το αν είναι εφικτό το πρόγραμμα Human Genome Initiative, ενώ το 1990 το DOE και NIH παρουσίασαν ένα πενταετές πλάνο για το πρόγραμμα. Παρόλα αυτά τότε υπολογίσθηκε πως το πρόγραμμα θα διαρκούσε 15 έτη και θα κόστιζε περίπου 3 εκατομμύρια δολάρια.

Το πρόγραμμα US Human Genome Initiative εγκαθίδρυσε την αλληλούχιση ως ένα σημαντικό επίτευγμα της ιατρικής και βιολογίας καθώς παράλληλα αλληλουχίστηκαν τα γονιδιώματα και άλλων οργανισμών «μοντέλων» όπως είναι το βακτήριο *E.coli*, η ζύμη *S. cerevisiae*, το *C. elegans*, η δροσόφιλα *D. Melanogaster*, το ποντίκι *Mus domesticus* ακόμα και το φυτό *Arabidopsis thaliana*. Το κρατικό πρόγραμμα US Human Genome Initiative του NIH, ήταν στην ουσία μία παγκόσμια προσπάθεια στην οποία συμμετείχαν εργαστήρια από τις Ηνωμένες Πολιτείες, την Ευρώπη και την Ιαπωνία, ενώ κάθε ερευνητικό κέντρο εστίαζε σε μια συγκεκριμένη περιοχή του γονιδιώματος. Το 1994, δημοσιεύθηκε ένα λεπτομερέστατος γενετικός χάρτης του ανθρώπινου γονιδιώματος ο οποίος περιελάμβανε 5,840 γενετικούς τόπους (Murray et al., 1994), ενώ το 1999, έγινε δυνατή η ολοκληρωμένη αλληλούχιση του πρώτου ανθρώπινου χρωμοσώματος (Χρωμόσωμα: 22) (Dunham et al., 1999). Στις 25 Ιουνίου του 2000 ο πρόεδρος των Η.Π.Α Clinton με τον Πρωθυπουργό της Μ. Βρετανίας Tony Blair έκαναν διάγγελμα για την αλληλούχιση του ανθρώπινου γονιδιώματος που πραγματοποιήθηκε από το κρατικό πρόγραμμα και την ιδιωτική εταιρεία Celera, ανακοινώνοντας ότι αποτελούσε ορόσημο στην Ιατρική επιστήμη και την ανθρωπότητα, αναμένοντας σημαντικά αποτελέσματα από τις πληροφορίες που θα προσέφερε η αλληλούχιση του ανθρώπινου γονιδιώματος. Το 2001 οι αλληλουχίες του ανθρώπινου γονιδιώματος από το κρατικό πρόγραμμα και τη Celera δημοσιεύθηκαν την ίδια εβδομάδα στα περιοδικά Science και Nature, αντίστοιχα (Lander et al. 2001; Venter et al. 2001).



### 3.3 Μεταλλάξεις

#### 3.3.1 Γενικά στοιχεία

Οι ζώντες οργανισμοί χαρακτηρίζονται από τεράστια ποικιλομορφία καθώς κάθε άτομο διαφέρει γενετικά από το άλλο, όχι μόνο ανάμεσα στα διάφορα είδη αλλά και μέσα στο ίδιο είδος, ακόμα και στον ίδιο πληθυσμό και οικογένεια. Η ποικιλομορφία αυτή αποτελεί προϋπόθεση για την εξέλιξη, δεδομένου ότι οι οργανισμοί που θα επιδείξουν μεγαλύτερη προσαρμοστικότητα στο περιβάλλον, θα παράγουν περισσότερους απογόνους και θα υπερισχύσουν στις επόμενες γενιές. Βάση της ποικιλομορφίας και, κατά συνέπεια, κινητήριος δύναμη για την εξέλιξη είναι οι μεταλλάξεις.

Με τον όρο «μετάλλαξη» χαρακτηρίζεται μία μόνιμη μεταβολή στο γενετικό υλικό ενός οργανισμού. Οι μεταλλάξεις συμβαίνουν με τυχαίο τρόπο μέσω διαφορετικών μηχανισμών, χωρίς αυτό να σημαίνει ότι δεν υπόκεινται και στην επίδραση του περιβάλλοντος. Πιθανές αιτίες εμφάνισης μίας μετάλλαξης αποτελούν λάθη κατά την αντιγραφή του DNA ή η επίδραση μεταλλαξιογόνων παραγόντων στο DNA. Το αποτέλεσμα μίας μετάλλαξης μπορεί να είναι η απώλεια λειτουργίας του γονιδίου ή η απόδοση σε αυτό μίας λειτουργίας διαφορετικής από αυτήν του φυσιολογικού γονιδίου. Μεταλλάξεις μπορούν να συμβούν σε οποιοδήποτε κύτταρο, γαμετικό ή σωματικό. Μόνο οι μεταλλάξεις σε κύτταρα γαμετικής σειράς (germline) μπορούν να μεταβιβαστούν από τη μία γενιά στην επόμενη, σε αντίθεση με τις μεταλλάξεις που συμβαίνουν σε σωματικά κύτταρα (somatic) οι οποίες δεν κληρονομούνται. Εκτός από τις μεταλλάξεις στην κωδικοποιούσα περιοχή ενός γονιδίου και στα όρια ιντρονίου/εξονίου, σημαντικό ρόλο παίζουν και οι χρωμοσωμικές μεταλλάξεις που επιδρούν σε ολόκληρο το χρωμόσωμα ή σε ένα μεγάλο τμήμα αυτού.

Υπάρχουν μεταλλάξεις σε επίπεδο γονιδίου οι οποίες επιγραμματικά είναι οι:

- Σιωπηρές παραλλαγές
- Παρανοηματικές Παραλλαγές (Missense Mutations)
- Μη νοηματικές Μεταλλάξεις (Nonsense Mutations)
- Μεταλλάξεις Αλλαγής του Πλαισίου Ανάγνωσης (Frameshift Mutations)
- Μεταλλάξεις στα όρια ιντρονίων/εξονίων
- Μεταλλάξεις στην 5' - μη κωδικοποιούσα περιοχή (5' - UTR).

Επίσης υπάρχουν και οι Γονιδιακές Αναδιατάξεις οι οποίες αναφέρονται στην απαλοιφή ή στον διπλασιασμό (ή ακόμα και τον τριπλασιασμό) μεγάλων περιοχών ενός γονιδίου στις οποίες περιλαμβάνονται ένα ή και περισσότερα εξώνια καθώς και ιντρονικές περιοχές.

Χαρακτηριστικό παράδειγμα μεταλλάξεων και εμφάνισης καρκίνου είναι τα ογκοκατασταλτικά γονίδια BRCA1 και BRCA2 που αποτελούν τα πιο σημαντικά γονίδια τα οποία σχετίζονται με υψηλή προδιάθεση για εμφάνιση καρκίνου του μαστού ή/και των ωοθηκών (Futreal et al., 1994). Το φάσμα των μεταλλάξεων των δύο αυτών γονιδίων, ποικίλει ανάλογα με τον πληθυσμό που μελετάται κάθε φορά και περιλαμβάνει τόσο σημειακές μεταλλάξεις και μικρές απαλοιφές και ενθέσεις βάσεων, οι οποίες αποτελούν και το πιο συχνό φαινόμενο απενεργοποίησης του γονιδίου, όσο και μεγάλες γονιδιακές αναδιατάξεις. Οι μεταλλάξεις αυτές, κατανέμονται σε όλο το μήκος του γονιδίου, αν και σε συγκεκριμένους πληθυσμούς, παρατηρείται συσσώρευση μεταλλάξεων σε ένα τμήμα του γονιδίου.

Η ανίχνευση των γονιδιακών αναδιατάξεων δε μπορεί να γίνει με τις κλασικές μεθόδους ανίχνευσης σημειακών μεταλλάξεων που βασίζονται στην τεχνική της PCR, αλλά απαιτούνται άλλες τεχνικές προσδιορισμού του αριθμού των αντιγράφων των γονιδίων. Οι κυριότερες τεχνικές που χρησιμοποιούνται είναι η τεχνική της Πολλαπλής και ποσοτικής αλυσιδωτής αντίδρασης της πολυμεράσης μικρών φθορίζόντων τμημάτων (QMPSF) και η τεχνική της Πολλαπλής ενίσχυσης τμημάτων DNA εξαρτώμενων από τη σύνδεση ιχνηθετών (MLPA), ενώ πρόσφατα η αλληλούχιση επόμενης γενιάς αποτελεί πλέον το πιο σημαντικό εργαλείο ανίχνευσης αυτών των γονιδιακών αναδιατάξεων (Desai and Jere, 2012).

## **3.4 Αλληλούχιση Επόμενης Γενιάς (Next Generation Sequencing)**

### **3.4.1 Γενικά στοιχεία**

Οι μεταγραφικοί παράγοντες και άλλες πρωτεΐνες οι οποίες αλληλεπιδρούν με τη χρωματίνη, είναι ουσιώδεις στη σύνδεση γονότυπου-φαινότυπου και ο προσδιορισμός της αλληλεπίδρασής τους με το DNA και τους μηχανισμούς ρύθμισης είναι κρίσιμος για τη διάσπαση της πολυπλοκότητας πολλών βιολογικών και παθολογικών καταστάσεων. Οι επιγενετικές πληροφορίες είναι συμπληρωματικές στις γονοτυπικές αναλύσεις και τις αναλύσεις γονιδιακής έκφρασης και διάφορες παραδοσιακές μέθοδοι έχουν επιτυχώς

αναγνωρίσει συγκεκριμένες τροποποιήσεις αλληλεπιδρώντων με το DNA πρωτεϊνών, καθώς επίσης και σημεία πρόσδεσης μεταγραφικών παραγόντων και τους ρόλους τους στη ρύθμιση συγκεκριμένων γονιδίων. Ο ρόλος των μεταγραφικών παραγόντων και της αλληλεπίδρασής τους με τη χρωματίνη είναι πολύ σημαντικός στην απορρύθμιση βιολογικών διαδικασιών που όχι σπάνια οδηγούν σε παθολογικές καταστάσεις, πολλές φορές εξαιρετικά επικίνδυνες όπως ο καρκίνος. Για παράδειγμα, έχειδειχθεί ότι ο μεταγραφικός παράγοντας Prox1 επάγει την πρόοδο του καρκίνου του παχέος εντέρου προάγοντας τη μετάβαση του φαινότυπου από καλοήγη σε δυσπλαστικό (Petrona et al., 2008). Ένα άλλο παράδειγμα είναι ο έλεγχος της ρύθμισης του μεταγραφικού παράγοντα Egr-1 από την κουρκουμίνη, η οποία καταστέλλει την έκφραση του υποδοχέα του επιδερμικού αυξητικού παράγοντα, έχοντας σαν αποτέλεσμα την αναστολή της ανάπτυξης ανθρώπινων καρκινικών κυττάρων παχέος εντέρου (Chen et al., 2006) καθώς επίσης και η δραστηριότητα του παράγοντα Cdx2 ο οποίος περιορίζει τον πολλαπλασιασμό ανθρώπινων καρκινικών κυττάρων παχέος εντέρου αναστέλλοντας τη μεταγραφική δραστηριότητα του διμερούς συμπλόκου β-κατενίνης-T-κυττάρου (Guo et al., 2010). Πέρα από τις δράσεις μεταγραφικών παραγόντων, άλλοι σημαντικοί επιγενετικοί παράγοντες συστηματικών διαταράξεων που κατευθύνουν νοσηρές καταστάσεις σε διάφορα είδη καρκίνων είναι οι μεθυλιώσεις του DNA (Das and Singal, 2004).

Τα τελευταία χρόνια, τεράστιες τεχνολογικές εξελίξεις έδωσαν στις επιστήμες της ζωής τα μέσα να αποκτούν δεδομένα υψηλής ποιότητας σε σχετικά μικρά χρονικά διαστήματα. Πιο συγκεκριμένα, με την εισαγωγή νέων πλατφορμών μαζικής παράλληλης αλληλούχισης (Roche 454 Genome Sequencer, Illumina Genome Analyzer, Applied Biosystems SOLiD και πρόσφατα η Helicos HeliScope) αναγγέλλοντας τη δεύτερη γενιά αλληλούχισης του DNA (*Αλληλούχιση Επόμενης Γενιάς* ή Next Generation Sequencing-NGS), οι ερευνητές είναι ικανοί να ανακτούν 50-400 εκατομμύρια μικρά κομμάτια αλληλουχιών (Πίνακας 1). Ένας δυνατός συνδυασμός τεχνικών ο οποίος έχει σαν αποτέλεσμα πολλαπλές βιολογικές εφαρμογές όπως γονιδιωματική αλληλούχιση εκ νέου ή επαναλληλούχιση, η μελέτη μεθυλιώσεων της χρωματίνης και πρωτεϊνικές αλληλεπιδράσεις σε επίπεδο ολόκληρου γονιδιώματος, είναι αυτός του NGS με την ανοσοκατακρήμνιση Χρωματίνης (Chromatin Immunoprecipitation -ChIP), μια ισχυρή μέθοδος που ανιχνεύει επιλεκτικά αλληλουχίες DNA στις οποίες έχει προσδεθεί μια συγκεκριμένη πρωτεΐνη σε ζωντανά κύτταρα, και της

βαθιάς αλληλούχισης (ChIP-Seq). Ένα τυπικό πείραμα ChIP αποτελείται από τη διασταυρωτή διασύνδεση πρωτεϊνικών συμπλόκων που έρχονται σε επαφή με το DNA στα σημεία πρόσδεσής τους. Έπειτα, η χρωματίνη κόβεται σε μικρά θραύσματα και η συγκεκριμένη περιοχή του DNA που αλληλεπιδρά με την πρωτεΐνη απομονώνεται με ανοσοκατακρήμνιση. Παράγεται μια ανάγνωση των σημείων πρόσδεσης της πρωτεΐνης σε επίπεδο ολόκληρου γονιδιώματος με την αλληλούχιση εκατομμυρίων διαφορετικών θραυσμάτων DNA (ChIP-Seq). Ο σκοπός της ανάλυσης δεδομένων ChIP-Seq είναι να βρεθούν εμπλουτισμένες γονιδιωματικές περιοχές σε μια λίμνη συγκεκριμένων κατακρημνισμένων θραυσμάτων DNA και το αποτέλεσμα μεθόδων και λογισμικού εύρεσης κορυφών (peaks) είναι συνήθως μια λίστα από κληθείσες κορυφές (peak calls) που αποτελούν τις γονιδιωματικές περιοχές που έχει προβλεφθεί ότι καταλαμβάνονται από την πρωτεΐνη (Valouev et al., 2008).

Τις τελευταίες τρεις δεκαετίες, η τεχνολογία αλληλούχισης DNA έχει κάνει μεγάλα άλματα. Οι μεθοδολογίες έχουν εξελιχθεί και από σήμανση με ραδιενέργεια περάσαμε σε σήμανση με φθορίζουσες ουσίες, και από ηλεκτροφόρηση σε πολυακρυλαμίδη σε ηλεκτροφόρηση επί τριχοειδούς (Tettelin and Feldblyum, 2009).

|                                   | <b>Roche (454):<br/>GS FLX Titanium</b> | <b>Illumina:<br/>Genome Analyzer II</b>               | <b>Applied Biosystems:<br/>SOLiD</b>                        |
|-----------------------------------|-----------------------------------------|-------------------------------------------------------|-------------------------------------------------------------|
| <b>Χημεία</b>                     | Πυροαλληλούχιση                         | Αλληλούχιση με σύνθεση βασισμένη στην στην πολυμεράση | Αλληλούχιση με βάση την την πρόσδεση μικρών μορίων (βάσεις) |
| <b>Προσέγγιση</b>                 | PCR σε μικρογαλακτώματα                 | Γεφυρωτή ενίσχυση                                     | PCR σε μικρογαλακτώματα                                     |
| <b>MB/χρήση</b>                   | 400-600 MB                              | 1300 MB                                               | 3000 MB                                                     |
| <b>Χρόνος/χρήση</b>               | 10 ώρες                                 | 4 μέρες                                               | 5 μέρες                                                     |
| <b>Μέγεθος μικρών αλληλουχιών</b> | 400 bp                                  | έως 75 bp                                             | 35 bp                                                       |

**Πίνακας 1: Σύγκριση των πλατφορμών NGS της Roche, της Illumina, και της Applied Biosystems**

Τις αλλαγές τις καθοδηγεί η ανάγκη για περισσότερη, γρηγορότερη και φθηνότερη αλληλούχιση DNA. Έτσι έχουν αναπτυχθεί πλατφόρμες αλληλούχισης με υψηλή απόδοση (high throughput) όπου η ταχύτητα είναι μεγάλη αλλά και το κόστος

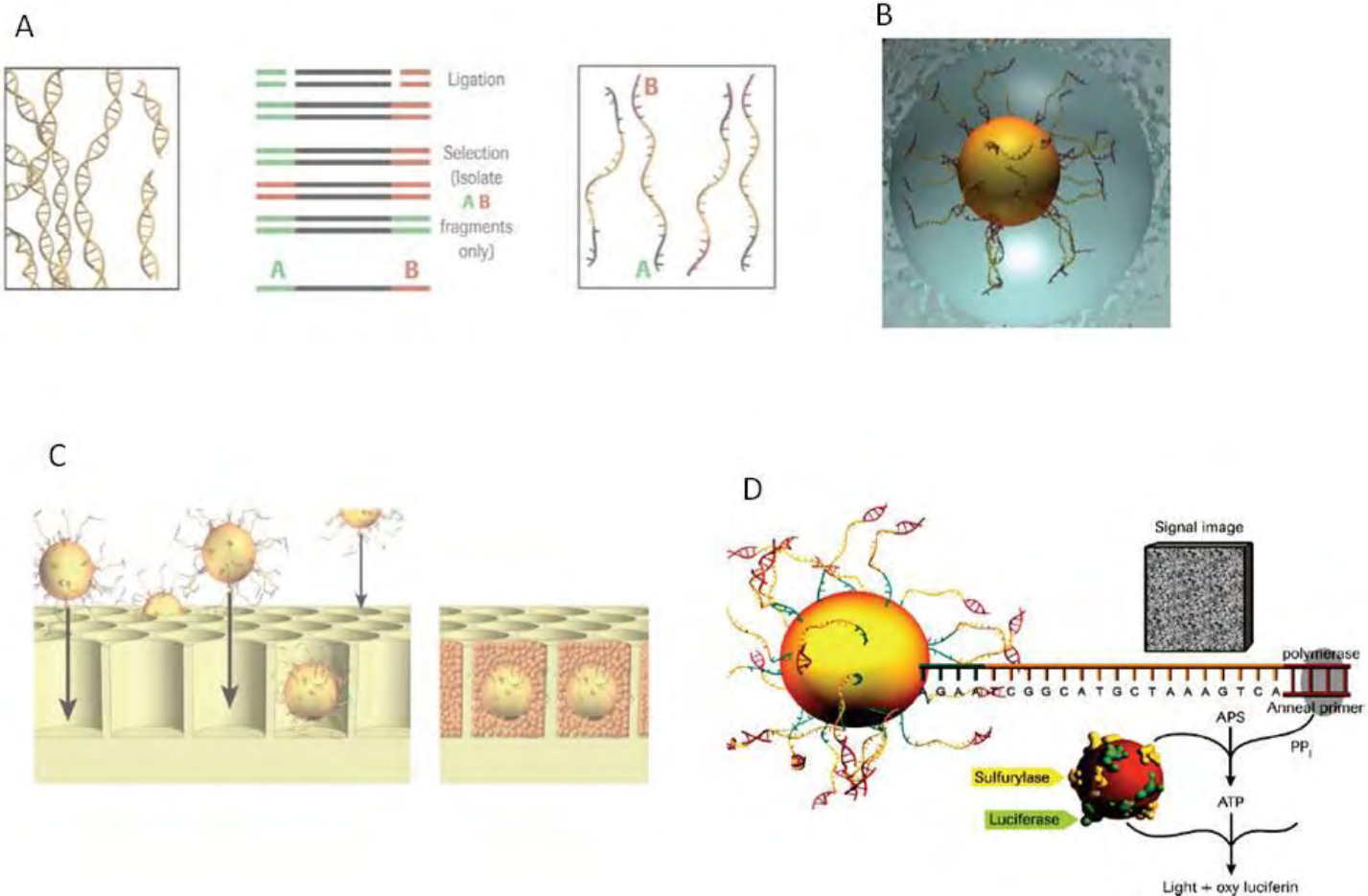
σχετικά με τον αριθμό βάσεων που αλληλουχούν σε σύγκριση με την κλασσική μέθοδο αλληλούχισης (Sanger). Μέχρι στιγμής τρεις πλατφόρμες next generation ξεχωρίζουν, η *G5 FLX (454)*, η *Illumina (Solexa)* και η *ABSOLiD*.

### **3.4.2 Genome Sequencer FLX System (454)**

Η πλατφόρμα της *Roche GS FLX*, βασίζεται στην τεχνολογία αλληλούχισης «κατά την σύνθεση» (pyrosequencing) που αναπτύχθηκε από την 454 Life Sciences και ήταν η πρώτη διαθέσιμη πλατφόρμα «next generation» αλληλούχισης (Margulies et al., 2005). Το δείγμα DNA θραυσματοποιείται σε μικρά τμήματα. Δύο μικροί αντάπτορες (A&B) συνδέονται στα θραύσματα. Οι αντάπτορες περιέχουν αλληλουχίες για ενίσχυση και αλληλούχιση καθώς και μια «ακολουθία κλειδί». Ο αντάπτορας B συνδέεται στο 5' άκρο του DNA με βιοτίνη ώστε να ακινητοποιηθούν τα θραύσματα σε μαγνητικά σφαιρίδια (beads) με επικάλυψη στρεπταβιδίνης. Τα δίκλινα θραύσματα δένουν στα σφαιρίδια και στην συνέχεια αποδιατάσσονται ώστε να ελευθερωθεί η συμπληρωματική ακολουθία η οποία περιέχει τον A και τον B αντάπτορα. Αυτές οι ακολουθίες αποτελούν την μητρική μονόκλωνη DNA (ssDNA) βιβλιοθήκη. Για την ενίσχυση τους στο GS FLX, χρησιμοποιείται η τεχνική της γαλακτωματοποίησης (emulsion-based clonal amplification, emPCR) (Dressman et al., 2003). Η ssDNA βιβλιοθήκη στη συνέχεια υβριδίζεται σε άλλα σφαιρίδια τα οποία έχουν επικάλυψη με ακινητοποιημένους εκκινητές. Η διαδικασία αυτή έχει βελτιστοποιηθεί ώστε να παράγονται σφαιρίδια που στο καθένα να έχει υβριδιστεί μόνο ένα θραύσμα. Στη συνέχεια τα σφαιρίδια μαζί με την βιβλιοθήκη, γαλακτωματοποιούνται μαζί με τα αντιδραστήρια που απαιτούνται για την ενίσχυση σε ένα διάλυμα (water-in-oil). Κάθε σφαιρίδιο φέρει ένα θραύσμα της βιβλιοθήκης και παγιδεύεται σε ένα γαλάκτωμα όπου λαμβάνει χώρα μια ανεξάρτητη αντίδραση PCR. Μετά την ενίσχυση, τα σφαιρίδια ελευθερώνονται και το κάθε ένα φέρει πολλά αντίγραφα του ίδιο αρχικού ssDNA. Για την αλληλούχιση, τα σφαιρίδια τοποθετούνται σε μια πλάκα που φέρει πηγάδια (PicoTiterPlate), που περιέχουν τα απαραίτητα ένζυμα καθώς και σφαιρίδια επιστοίβασης (Εικόνα 3).

Σε ρευστή μορφή προστίθενται τα νουκλεοτίδια, ενός τύπου κάθε φορά και με συγκεκριμένη σειρά. Όταν ένα νουκλεοτίδιο είναι συμπληρωματικό με την μήτρα, ενσωματώνεται στην νεοσυντιθέμενη αλυσίδα από την πολυμεράση. Η ενσωμάτωση συνοδεύεται από ισομοριακή απελευθέρωση ενός πυροφωσφορικού.

Ένα μειονέκτημα του pyrosequencing είναι η αδυναμία αλληλούχισης ομοπολυμερών ακολουθιών όπως τα polyAs καθώς και το κόστος ανά βάση είναι υψηλότερο σε σχέση με τις άλλες δύο πλατφόρμες (SOLiD και Solexa) (Rothberg and Leamon, 2008). Το μεγάλο πλεονέκτημα του pyrosequencing έναντι των άλλων δύο είναι το μέγεθος των διαβασμάτων (Metzker, 2009). Σε 10 ώρες μπορούν να αλληλουχηθούν 1.000.000 ανεξάρτητα μόρια DNA μέσου μεγέθους 400bp (0.4 Kb).



**Εικόνα 3. Σύνοψη της λειτουργίας του Roche/454**

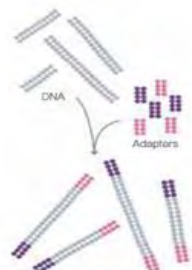
- A. Προετοιμασία της ss cDNA βιβλιοθήκης,
- B. Μοναδιαία ενίσχυση των τμημάτων σε γαλάκτωμα,
- C. Εναπόθεση των σφαιριδίων στην πλάκα PicoTiterPlate,
- D. Αλληλούχιση κατά την σύνθεση ([www.454.com](http://www.454.com))

### **3.4.3 The Illumina (Solexa) Genome Analyzer**

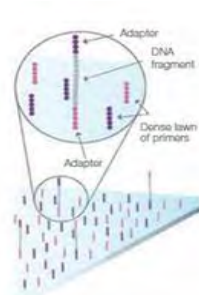
Η πλατφόρμα αλληλούχισης της Solexa, είναι εμπορικά διαθέσιμη από το 2006. Η αρχή

λειτουργίας βασίζεται στην αλληλούχιση «κατά την σύνθεση» (Εικόνα 4). Το DNA θραυσατοποιείται σε τμήματα μικρότερα των 80 βάσεων. Τα θραύσματα αποκτούν αμβλεία άκρα, φωσφορυλιώνονται, και προστίθεται μια αδενίνη στα 3' άκρα τους. Έπειτα συνδέονται αντάπτορες με μία θυμίνη και ακολουθεί αποδιάταξη και υβριδοποίηση σε συμπληρωματικούς αντάπτορες που είναι ακινητοποιημένοι σε μια στερεή επιφάνεια. Κάθε μονόκλωνο τμήμα που είναι ακινητοποιημένο στο ένα άκρο, δημιουργεί μια δομή «γέφυρας» λόγω υβριδισμού του άλλου άκρου του με κάποιον γειτονικό ακινητοποιημένο αντάπτορα. Προστίθενται τα απαραίτητα αντιδραστήρια και ακολουθεί ενίσχυση με «bridge PCR» αφού οι αντάπτορες λειτουργούν ως εκκινήτες (Adessi *et al.*, 2000; Fedurco *et al.*, 2006 ). Μετά από αρκετούς κύκλους, έχουμε περίπου 1000 αντίγραφα του αρχικού ssDNA δημιουργώντας ένα cluster. Ακολουθεί η σύνθεση DNA και η αλληλούχιση με προσθήκη των τεσσάρων νουκλεοτιδίων «ανάστροφου τερματισμού» που το καθένα έχει σημανθεί με διαφορετική φθορίζουσα χρωστική (Εικόνα 4). Κατά την σύνθεση της συμπληρωματικής αλυσίδας, ανιχνεύεται από μια κάμερα CCD ο τύπος του νουκλεοτιδίου που ενσωματώθηκε (λόγω φθορισμού) καθώς και η θέση του πάνω στην επιφάνεια. Στην συνέχεια αφαιρείται από το 3' άκρο της βάσης η ομάδα τερματισμού και η φθορίζουσα χρωστική και ακολουθεί άλλος ένας κύκλος. Έτσι συνθέτοντας την συμπληρωματική ακολουθία και ανιχνεύοντας βάση-βάση τον φθορισμό, γίνεται αλληλούχιση του cluster και κατ'επέκταση του αρχικού ssDNA.

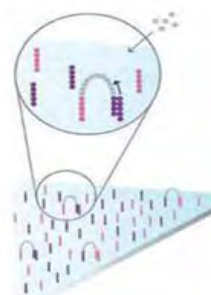
Το 2008 έγινε μια αναβάθμιση (Genome Analyzer II) και πλέον μετά την ολοκλήρωση της αλληλούχισης, οι μητρικές αλυσίδες αναγεννούνται *in situ* και ακολουθεί δεύτερος γύρος αλληλούχισης από το άλλο άκρο. Αρχικά τα τμήματα που αλληλουχήθηκαν αφαιρούνται, γίνεται «bridge PCR» στις συμπληρωματικές ακολουθίες και ακολουθεί ξανά αλληλούχιση (paired-end sequencing). Έτσι μπορούν να αλληλουχηθούν τμήματα μεγέθους 2x100 bp από θραύσματα μεγέθους 200bp έως 5kb και σε κάθε «τρέξιμο» το Genome Analyzer II αλληλουχεί 45-50 Gb. Σε σύγκριση με την κλασσική μέθοδο αλληλούχισης (Sanger), η πλατφόρμα της Illumina, παράγει τρομακτικά περισσότερα αποτελέσματα με πολύ λιγότερο κόστος ανά βάση και σε λιγότερο χρόνο αλλά η πιθανότητα λάθους είναι μεγαλύτερη (Metzker, 2009). Τα λάθη κατά την αλληλούχιση μπορούν να διορθωθούν εάν υπάρχει αρκετή επικάλυψη μεταξύ των διαβασμάτων ωστόσο το μικρό τους μέγεθος το καθιστά δύσκολο.



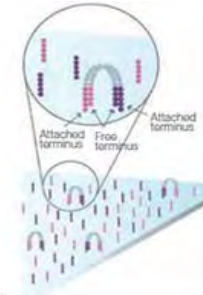
**Figure 1**  
Randomly fragment genomic DNA and ligate adaptors to both ends of the fragments.



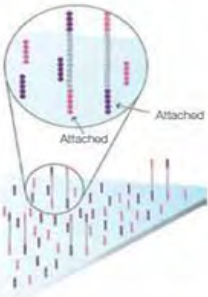
**Figure 2**  
Bind single-stranded fragments randomly to the inside surface of the flow cell channels.



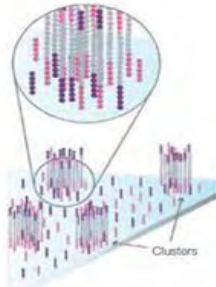
**Figure 3**  
Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.



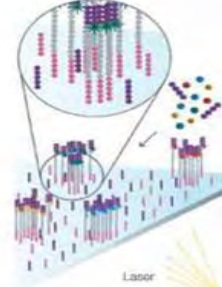
**Figure 4**  
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.



**Figure 5**  
Denaturation leaves single-stranded templates anchored to the substrate.



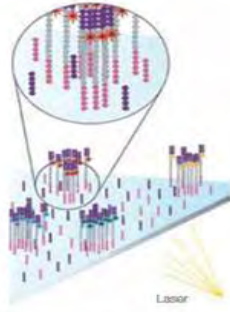
**Figure 6**  
Several million dense clusters of double-stranded DNA are generated in each channel of the flow cell.



**Figure 7**  
The first sequencing cycle begins by adding four labeled reversible terminators, primers, and DNA polymerase.



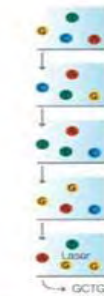
**Figure 8**  
After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.



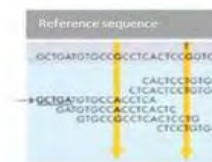
**Figure 9**  
The next cycle repeats the incorporation of four labeled reversible terminators, primers, and DNA polymerase.



**Figure 10**  
After laser excitation, the image is captured as before, and the identity of the second base is recorded.



**Figure 11**  
The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time.



**Figure 12**  
The data are aligned and compared to reference, and sequencing differences are identified.

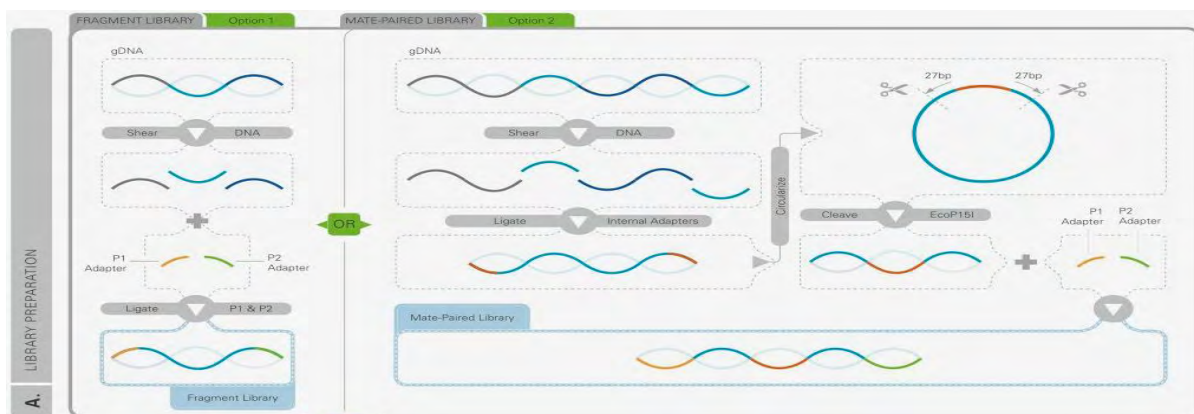
### Εικόνα 4. Σύνοψη της λειτουργίας της πλατφόρμας Illumina.

1. Προετοιμασία του δείγματος DNA,
2. Προσκόλληση των τμημάτων στην επιφάνεια,
3. Ενίσχυση με bridge PCR,
4. Δημιουργία δίκλωνων τμημάτων,
5. Αποδιάταξη αυτών,
6. Ενίσχυση,
7. Προσδιορισμός πρώτης βάσης,
8. Καταγραφή σήματος,
9. Προσδιορισμός δεύτερης βάσης,
10. Καταγραφή σήματος,
11. Αλληλούχιση μετά από πολλούς κύκλους,
12. Στοίχιση αποτελεσμάτων. ([www.illumina.com](http://www.illumina.com))



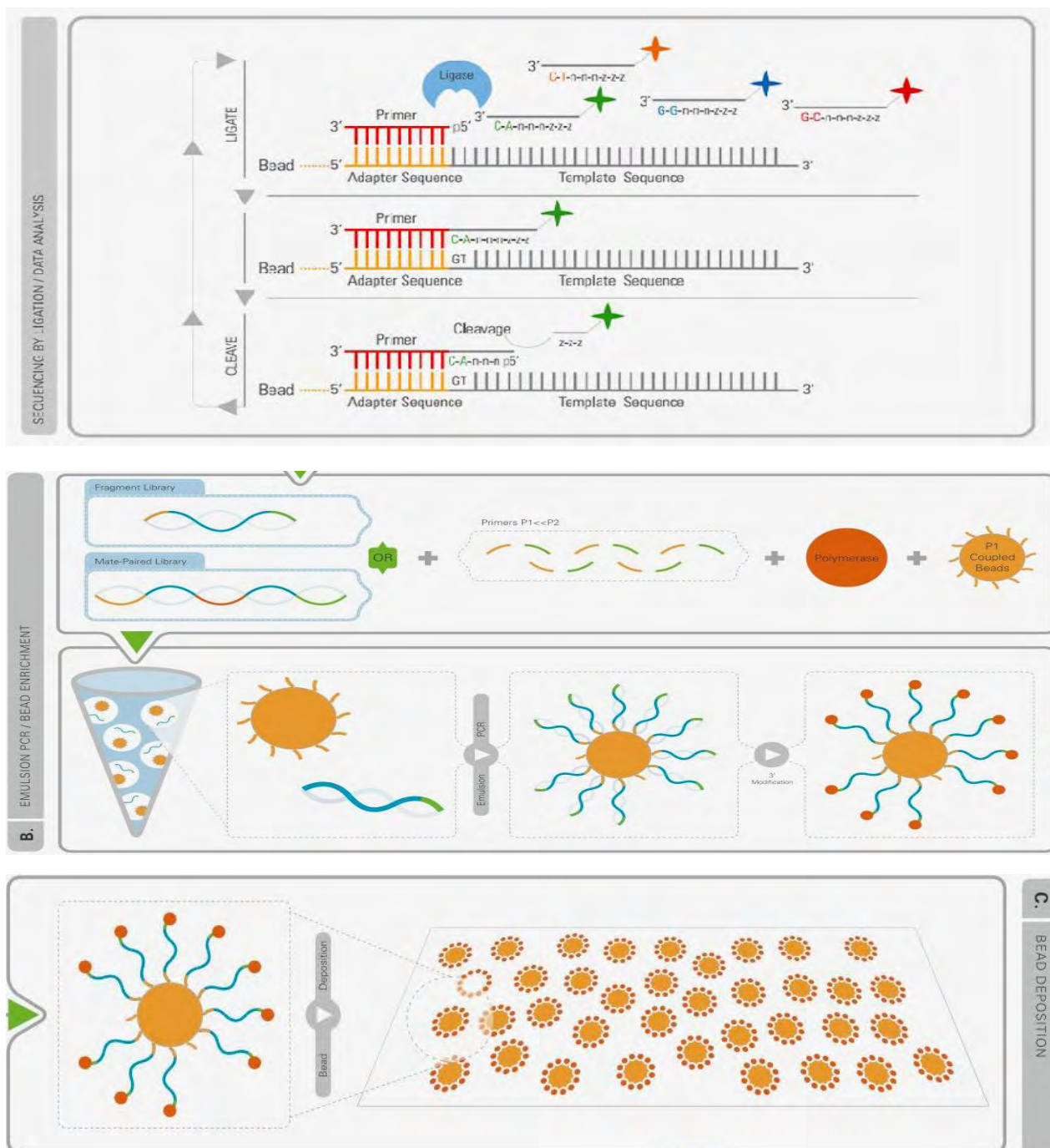
### **3.4.4 The Applied Biosystem SOLiD system**

Η πλατφόρμα ABSOLiD βασίζεται στην τεχνολογία αλληλούχισης «μέσω ολιγονουκλεοτιδικής σύνδεσης και ανίχνευσης». Εδώ υπάρχει η δυνατότητα δημιουργίας δύο ειδών βιβλιοθήκων, είτε θραυσμάτων είτε «συζευγμένων τμημάτων» (mate-paired) ανάλογα με τον ερευνητικό στόχο. Στην πρώτη περίπτωση εφαρμόζονται υπέρηχοι και προκύπτουν τμήματα μεγέθους 60-90 bp. Στην δεύτερη, με τον ίδιο τρόπο έχουμε μεγαλύτερα τμήματα (0,6-10 kb), επιλέγεται μετά από ηλεκτροφόρηση το κατάλληλο μέγεθος (ανάλογα την περίπτωση) και γίνεται σύνδεση με έναν συνθετικό εκκινητή (CAP adapter). Ακολουθεί αραίωση και συνθήκες που ευνοούν την ενδομοριακή κυκλοποίηση δημιουργώντας έτσι μια βιβλιοθήκη τυχαίων κυκλικών τμημάτων συγκεκριμένου μεγέθους. Στην συνέχεια γίνεται πέψη με μια περιοριστική ενδονουκλεάση τύπου III (EcoP15I, θέση αναγνώρισης στον CAP και κόβει 25 και 27 βάσεις μακριά). Έτσι προκύπτει μια βιβλιοθήκη από συζευγμένα τυχαία 27μερή τα οποία αρχικά στο γονιδίωμα διαχωρίζονταν σε τμήματα μεγέθους που επιλέχθηκε στην ηλεκτροφόρηση.



Τα θραύσματα DNA συνδέονται με αντάπτορες και ακινητοποιούνται σε σφαιρίδια όπου ενισχύονται με PCR. Μετά το PCR, οι μητρικές αλυσίδες αποδιατάσσονται, γίνεται επιλογή των σφαιριδίων που φέρουν τα ενισχυμένα τμήματα και γίνεται μια τροποποίηση στο 3' άκρο ώστε να είναι εφικτή η ακινητοποίηση τους σε γυάλινη επιφάνεια με ομοιοπολικό δεσμό. Η πρωτοτυπία του συστήματος είναι η χρήση της DNA λιγάσης και σημασμένων ολιγονουκλεοτιδίων για την αλληλούχιση (Housby and Southern, 1998). Στην αρχή, υβριδίζεται ένας εκκινητής στον αντάπτορα που έχει προστεθεί στο DNA. Στην συνέχεια προστίθεται ένα pool οκταμερών με όλους τους πιθανούς συνδυασμούς βάσεων (A, C, G και T) στις θέσεις 1 έως 5 (1024 συνδυασμοί).

Μόνο το ολιγονουκλεοτίδιο του οποίου οι 5 πρώτες βάσεις είναι συμπληρωματικές με την άγνωστη αλυσίδα θα είναι στην κατάλληλη θέση ώστε να δημιουργηθεί 3'-5' φωσφοδιεστερικός δεσμός με τον εκκινητή. Έχουμε τέσσερις φθορίζουσες χρωστικές και κάθε οκταμερές είναι σημασμένο με μία, σε αντιστοιχία με τα τέσσερα πιθανά δίνουκλεοτίδια των θέσεων 1 και 2.



**Εικόνα 5. Αλληλούχηση κατά τη σύνδεση με ολιγονουκλεοτίδια με την πλατφόρμα ABSOLiD.**

Αρχικά πραγματοποιείται προετοιμασία της βιβλιοθήκης και στη συνέχεια η πρόσδεση και ενίσχυση σε γαλάκτωμα. Ακολουθεί η εναπόθεση των σφαιριδίων και η αλληλούχηση μέσω σύνδεσης. Αργότερα αλλάζει ο εκκινητής και η διαδικασία επαναλαμβάνεται ([www.appliedbiosystems.com](http://www.appliedbiosystems.com))

Μετά την ανίχνευση σήματος, αφαιρούνται οι βάσεις 6-8 και ακολουθεί νέος γύρος. Έπειτα από 7 κύκλους σύνδεσης έχουμε πιθανούς συνδυασμούς δινουκλεοτιδίων για τις θέσεις 4+5, 9+10, 14+15, 19+20, 24+25, 29+30 και 34+35. Προκειμένου να ολοκληρωθεί η αλληλούχιση, αφαιρούνται όλα τα προϊόντα σύνδεσης από την μήτρα και επαναλαμβάνεται από την αρχή η διαδικασία με εκκινητή που διαφέρει κατά 1, 2, 3 ή 4 βάσεις από τον προηγούμενο. Συνολικά γίνονται 5 φορές 7 κύκλοι σύνδεσης ολιγονουκλεοτιδίων ώστε να αλληλουχηθούν 25-35 βάσεις. Η πλατφόρμα SOLiD 3 Plus μπορεί να αλληλουχήσει πάνω από 60 Gb, το κόστος εξοπλισμού είναι μικρότερο από τις άλλες δύο πλατφόρμες ωστόσο το μικρό μέγεθος των διαβασμάτων περιορίζει τις εφαρμογές της. (Metzker, 2009). Μια διαγραμματική απεικόνιση της αλληλούχισης με την πλατφόρμα ABSOLiD παρουσιάζεται στην Εικόνα 5.

## 3.5 Εφαρμογές

### 3.5.1 Γενικά στοιχεία

Οι νέες τεχνολογίες αλληλούχισης έχουν ήδη εφαρμοστεί σε αρκετούς τομείς. Οι σημαντικότεροι είναι:

- de novo αλληλούχιση γονιδιωμάτων, πιο στοχευμένη επαναλληλούχιση για εντοπισμό μεταλλάξεων ή πολυμορφισμών
- αλληλούχιση μεταγράφων (deep sequencing) απευθείας από cDNAs χωρίς την ανάγκη βιβλιοθηκών σε βακτήρια, BACs κτλ
- Μεγάλης κλίμακας ανίχνευση μεθυλιώσεων σε DNA που έχει υποστεί χειρισμό με bisulfate
- Προσδιορισμός αλληλεπιδράσεων DNA-πρωτεΐνης μετά από ανοσοκατακρήμνιση χρωματίνης

Οι νέες αυτές τεχνολογίες έφεραν επανάσταση στην γονιδιωματική κάνοντας εφικτή την μαζική παράλληλη αλληλούχιση με χαμηλό κόστος, στο μέλλον όμως ίσως αναλάβουν ακόμα πιο νέες τεχνολογίες (Next-Next Generation Sequencing) τον ρόλο αυτό, αλληλουχώντας μοναδιαία μόρια (single-molecule sequencing). Προσπάθειες για να δημιουργηθούν τέτοιες πλατφόρμες έχουν ήδη γίνει. Η αλληλούχιση μοναδιαίων μορίων βοηθάει στο να αυξηθεί ο αριθμός τμημάτων DNA που μπορούν ανεξάρτητα να

αναλυθούν σε μια δεδομένη επιφάνεια επιτυγχάνοντας πολύ μεγαλύτερη αναλυτικότητα. Επίσης παρακάμπτεται το βήμα της ενίσχυσης σε clusters, μειώνοντας έτσι το κόστος. Πρώτη η Helicos Biosciences προσέφερε τέτοιες υπηρεσίες με το HeliScope το οποίο χρησιμοποιεί μια μικροτεχνική που βασίζεται στον φθορισμό και ονομάζεται TIRM (Total Internal Reflection Microscopy) για να ανιχνεύσει τον φθορισμό κατά την ενσωμάτωση νουκλεοτιδίων (Harris et al. 2008). Προς αυτή την λογική αλληλούχισης (single molecule) κινείται και η visiGen Biotechnologies, η οποία έχει κατασκευάσει μια νανοσυσκευή που καταγράφει τα νουκλεοτίδια που ενσωματώνει η DNA πολυμεράση σε πραγματικό χρόνο βασιζόμενη στην μεταφορά ενεργείας φθορισμού μέσω συντονισμού (FRET- Fluorescence Resonance Energy Transfer) (Blow, 2008). Τέλος η Pacific Biosciences έχει αναπτύξει την τεχνολογία SMRT (Single-Molecule Real-Time) κατά την οποία γίνεται αλληλούχιση «κατά την σύνθεση» μοναδιαίων μορίων και βασίζεται σε μια νανοδομή (ZMW, Zero- Mode Waveguide) όπου καταγράφεται πολυμερισμός σε πραγματικό χρόνο (Eid et al., 2009).

### **3.5.2 Αλληλούχιση επόμενης γενιάς και συγγενείς διαταραχές γλυκοζυλίωσης**

Οι συγγενείς διαταραχές γλυκοζυλίωσης αποτελούν μια ομάδα διαταραχών οι οποίες εμφανίζουν δυσμορφικές καταστάσεις, καθυστερημένη ανάπτυξη, επιληπτικές κρίσεις, καρδιακές δυσλειτουργίες, ηπατική ίνωση, αμφιβληστροειδοπάθεια και διαταραχές πήξης του αίματος (Jones et al., 2011). Περίπου το 50% των πρωτεϊνών του ανθρώπινου οργανισμού N-γλυκοζιώνονται, κάτι το οποίο αποτυπώνεται στην κλινική εμφάνιση αυτών των διαταραχών.

Περισσότεροι από 30 τύποι συγγενών διαταραχών γλυκοζυλίωσης έχουν περιγραφεί και προέρχονται από ελαττωματική λειτουργία γονιδίων. Ορισμένες μέθοδοι όπως η υγρή χρωματογραφία υψηλής πίεσης (High Performance Liquid Chromatography - HPLC) και η MALDI-TOF-MS (Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry) είναι ικανές να αναγνωρίσουν τα επίπεδα των ολιγοσακχαριτών και τη δομή των γλυκανών. Μια αρχική διάγνωση μπορεί να επιτευχθεί μέσω των παραπάνω βιοχημικών αναλύσεων και στη συνέχεια να ακολουθήσει αλληλούχιση συγκεκριμένων γονιδίων με την κλασικές μεθόδους. Με τον τρόπο αυτό όμως μόνο στο 40% των περιπτώσεων υπάρχει επιτυχής διάγνωση, ενώ με

την τεχνική NGS μπορούν να ανιχνευθούν κατά 100% οι μεταλλάξεις των γονιδίων που προκαλούν τη συγγενή διαταραχή γλυκοζυλίωσης (Jones et al., 2011). Επιπρόσθετα, υπάρχει η δυνατότητα μέσω αλγορίθμου που περιλαμβάνει και άλλα βιοχημικά και μοριακά τεστ, η διάγνωση που θα προκύψει να είναι σημαντικά βελτιωμένη. Τα κύρια γονίδια που ελέγχονται με την τεχνική NGS είναι 24 (*PMM2, ALG1, ALG2, ALG3, ALG6, ALG8, ALG9, ATP6V0A2, B4GALT1, COG1, COG7, COG8, DOLK, DPAGT1, DPM1, GNE, MGAT2, MOGS, MPDUI, MP, RFT1, SLC35A1, SLC35C1* και *TUSC3*) και έτσι πολύ γρήγορα και με μικρό κόστος μπορούν να ανιχνευθούν οι τυχόν μεταλλάξεις και να γίνει η ορθή διάγνωση και να παρασχεθεί η κατάλληλη θεραπεία. Επίσης, μέσω της τεχνικής NGS γίνεται εφικτή η ορθή διάγνωση σε ασθενείς οι οποίοι δεν παρουσιάζουν τα κλασσικά συμπτώματα των συγγενών διαταραχών γλυκοζυλίωσης (Jones et al., 2011).

### **3.5.3 Αλληλούχιση επόμενης γενιάς και καρκίνος**

Το ανθρώπινο γονιδίωμα αποτελείται από 3 εκατομμύρια βάσεις και είναι χωρισμένο σε 24 χρωμοσώματα, ενώ μόνο το 5% του γονιδιώματος περιέχει γονίδια από τα οποία το 1-2% είναι λειτουργικά ως προς την παραγωγή πρωτεϊνών (Lander et al., 2001; Venter et al, 2001). Το γονιδίωμα κάθε ανθρώπου είναι μοναδικό περιέχοντας μοναδιαίους πολυμορφισμούς νουκλεοτιδίων (Single Nucleotide Polymorphisms - SNPs). Οι πιο συχνές αλλαγές μεταξύ των γονιδιωμάτων διαφορετικών ανθρώπων είναι οι μονονουκλεοτιδικοί πολυμορφισμοί-SNPs, ενώ έχουν καταγραφεί παραπάνω από 38 εκατομμύρια τέτοιων αλλαγών (Acland et al., 2013)

[http://www.ncbi.nlm.nih.gov/projects/SNP/snp\\_summary.cgi](http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi)). Επιπρόσθετα, ένας μεγάλος αριθμός στην ποικιλομορφία των ομόλογων γονιδίων έχει ταυτοποιηθεί (Levy et al., 2007; Zogoroulos et al., 2007) και αναφέρεται ως ένας πολύ κρίσιμος παράγοντας στην ανάπτυξη των διάφορων μορφών καρκίνου (Stratton et al., 2009).

Οι γενετικές παραλλαγές που υπάρχουν στα γαμετικά κύτταρα κληρονομούνται από τους γονείς, ενώ οι γενετικές παραλλαγές που αποκτήθηκαν από ένα άτομο κατά τη διάρκεια της ζωής τους ονομάζονται σωματικές παραλλαγές. Η πλειονότητα των γονεϊκών γενετικών παραλλαγών οδηγούν σε διαφορές, όπως το χρώμα των μαλλιών και των ματιών, το ύψος και το χρώμα του δέρματος. Ωστόσο, μερικές γενετικές παραλλαγές μπορούν να συμβάλουν στην αυξημένη πιθανότητα εμφάνισης ασθενειών,

όπως είναι ο διαβήτης (Janipalli et al., 2012) η υπέρταση (Gong et al., 2012), ο αυτισμός (Casey et al., 2012), και ο καρκίνος (Link et al., 2011; Beroukhim et al., 2010).

| <i>Molecular marker</i>             | <i>Drug</i>                            | <i>Tumor site</i>       | <i>Guidelines and Recommendations</i> | <i>FDA label</i> |
|-------------------------------------|----------------------------------------|-------------------------|---------------------------------------|------------------|
| Estrogen and Progesterone receptors | Tamoxifen                              | Breast                  | NCCN, ASCO                            | Yes              |
| HER2/ERBB2                          | Trastuzumab                            | Breast                  | NCCN, ASCO, NICE                      | Yes              |
| KRAS                                | Cetuximab and Panitumumab              | Colon                   | NCCN, ASCO, EGAPP                     | Yes              |
| cKit                                | Imatinib                               | Gastrointestinal stroma | NA                                    | Yes              |
| ALK                                 | Crizotinib                             | Lung                    | NCCN                                  | Yes              |
| EGFR                                | Erlotinib and Afatinib                 | Lung                    | NCCN                                  | Yes              |
| BRAF                                | Vemurafenib, Trametinib and Dabrafenib | Melanoma                | NA                                    | Yes              |

ASCO – American Society of Clinical Oncology  
 CPIC – Clinical Pharmacogenetics Implementation Consortium  
 EGAPP – Evaluation of Genomic Applications in Practice and Prevention  
 EMA – European Medicines Agency  
 FDA – US Food and Drug Administration  
 NCCN – National Comprehensive Cancer Network

**Πίνακας 2. Παραδείγματα τυπικών μοριακών δεικτών σε σχέση με φαρμακογονιδιακές εφαρμογές για την αντιμετώπιση συγκεκριμένων ειδών καρκίνου. (Marrone et al., 2014)**

Οι παραδοσιακές φαρμακευτικές και γενετικές εφαρμογές για τον μοριακό χαρακτηρισμό του καρκίνου, βασίζονται στον έλεγχο ενός ή λίγων γονιδιακών δεικτών από τα κύτταρα του όγκου. Αυτές οι κλινικές εφαρμογές είναι καθιερωμένες και χρησιμοποιούνται προκειμένου να χαρακτηριστεί το είδος του όγκου κυρίως στους καρκίνους του μαστού, του παχέως εντέρου και του πνεύμονα και να επιλεγεί η κατάλληλη θεραπεία στοχευμένη σε συγκεκριμένους γονιδιακούς τόπους και SNP's (Πίνακας 2). Τα τελευταία χρόνια επειδή οι high-throughput τεχνολογίες των -omics (genomics, transcriptomics, proteomics) έχουν εξελιχθεί ιδιαίτερα με σημαντική πτώση του κόστους, οι ερευνητές δεν ελέγχουν πλέον μόνο ένα ή λίγα γονίδια, αλλά ένα πάνελ πολλών γονιδίων με αλληλούχιση του γονιδιώματος των καρκινικών κυττάρων. Όπως έχει προαναφερθεί οι τεχνολογίες NGS παρέχουν μια άμεση πληροφόρηση για τις γενετικές παραλλαγές που υπάρχουν στο γονιδίωμα και δεδομένου ότι πολλές φορές ο

καρκίνος προκύπτει από νέες σωματικές μεταλλάξεις, η τεχνολογία NGS αποτελεί ένα πολύ χρήσιμο εργαλείο για τη μελέτη των τροποποιήσεων αυτών. Με αυτό τον τρόπο, οι επιστήμονες θα κατανοήσουν σε βάθος τα μοριακά μονοπάτια που οδηγούν στην ογκογένεση, αλλά και θα μπορούν να χαρακτηρίσουν αποδοτικότερα τους καρκινικούς όγκους, ώστε στη συνέχεια να ακολουθήσει εξατομικευμένη θεραπεία με περισσότερες πιθανότητες επιτυχίας (MacConaill, 2013). Χαρακτηριστικό παράδειγμα στη μελέτη των Link et al., (2011), όπου χρησιμοποιήθηκε η τεχνική NGS ώστε να αποκωδικοποιηθεί πλήρως όλο το γονιδίωμα μιας ασθενούς που ανέπτυξε καρκίνο του μαστού, των ωοθηκών, αλλά και οξεία μυελογενή λευχαιμία. Η ασθενής δεν είχε κανένα οικογενειακό ιστορικό εμφάνισης καρκίνου και η εξέταση για πιθανές μεταλλάξεις στα γονίδια *BRCA1* και *BRCA2* ήταν αρνητική. Εφαρμόζοντας την τεχνική NGS αλληλουχήθηκε ολόκληρο το γονιδίωμα από κύτταρα του δέρματος (φυσιολογικά) και κύτταρα του μυελού των οστών (καρκινικά). Τα αποτελέσματα έδειξαν μια έλλειψη 3000 βάσεων στα εξώνια 7-9 του γονιδίου *TP53* (Tumor Protein 53) (από τα φυσιολογικά κύτταρα) και μια αλλαγή ολόκληρης περιοχής 18 μεγαβάσεων στο χρωμόσωμα 17 (από τα καρκινικά κύτταρα). Η έλλειψη των εξονίων 7-9 του γονιδίου *TP53* έχει ως επακόλουθο την έλλειψη του τμήματος πρόσδεσης σε DNA (DNA binding domain) με αποτέλεσμα μια δυσλειτουργική πρωτεΐνη. Οι συγγραφείς της παραπάνω μελέτης υποστήριξαν πως αυτή η έλλειψη ήταν υπεύθυνη για την εμφάνιση του καρκίνου στην ασθενή. Παρόμοια, οι Berger et al., (2011), εφάρμοσαν ανάλυση και αλληλούχιση ολόκληρου του γονιδιώματος (whole genome sequencing) από καρκινικούς προστατικούς όγκους και το αντιστοίχισαν με το φυσιολογικό γονιδίωμα, περιγράφοντας ένα νέο πάνελ γονιδίων που σχετίζεται με τον καρκίνο του προστάτη.

Ο Γονιδιακός Άτλας του Καρκίνου (The Cancer Genome Atlas – TCGA) είναι μια βάση δεδομένων στην οποία έχουν κατατεθεί όλες οι γονιδιακές μεταλλάξεις που συνδέονται με βιολογικά μονοπάτια πολλαπλασιασμού των κυττάρων, της απόπτωσης, του μεταβολισμού των καρκινικών όγκων και αλλαγών στη χρωματίνη των κυττάρων. Πλέον, η σύγχρονη ογκολογική πρακτική η οποία υπερτονίζει τη θέση και την ιστολογία του όγκου διανύει μια αλλαγή προς τη “γονιδιωματική ογκολογία” (Garraway et al., 2013).

Οι διαθέσιμες κλινικές εξετάσεις που περιλαμβάνουν την τεχνολογία αλληλούχισης επόμενης γενιάς (NGS), χρησιμοποιούνται για τον χαρακτηρισμό του γονιδιώματος του καρκίνου μέσω αλληλούχισης προκαθορισμένων γονιδίων των

οποίων οι μεταλλάξεις θα δώσουν σημαντικές πληροφορίες για το είδος του καρκίνου. Χρησιμοποιώντας μόνο ένα τεστ ανίχνευσης όλων των γονιδιακών αλλαγών από δείγμα βιοψίας, γίνεται εφικτή η αποδοτικότερη εξατομικευμένη θεραπεία ανάλογα με το είδος του καρκινικού όγκου (Garraway, 2013). Μάλιστα, η “γονιδιωματική ογκολογία” (genomics driven oncology) χρησιμοποιώντας πάνελ πολλών γονιδίων, δύναται να επεκτείνει το οπλοστάσιο των ιατρών ως προς τη θεραπεία ασθενειών οι οποίες δεν αντιμετωπίζονται με τις τυπικές θεραπείες, όπως είναι και οι μεταστατικές νόσοι. Παρακάτω θα περιγραφεί η εφαρμογή της τεχνικής NGS στην αντιμετώπιση του καρκίνου του μαστού και των ωθηκών, ως ένα πολύ χαρακτηριστικό παράδειγμα κλινικής εφαρμογής.

Περίπου το 30% των περιστατικών οικογενούς καρκίνου του μαστού σχετίζονται με μεταλλάξεις σε γονίδια προδιάθεσης με υψηλή διεισδυτικότητα (BRCA1, BRCA2, BRCAX) και αποτελούν τον κληρονομικό καρκίνο του μαστού (hereditary breast cancer). Ένα μικρό μέρος του κληρονομικού καρκίνου του μαστού συνδέεται με μεταλλάξεις και σε άλλα γονίδια όπως είναι τα TP53, STK11/LKB1 και PTEN. Η κλωνοποίηση των γονιδίων προδιάθεσης για τον κληρονομικό καρκίνο του μαστού και των ωθηκών BRCA1 και BRCA2 ήταν ένα σημαντικό βήμα για τη μελέτη της επίδρασής τους στις ασθένειες αυτές. Έχει βρεθεί ότι 15-20% των γυναικών με βεβαρημένο οικογενειακό ιστορικό καρκίνου του μαστού και 60-80% με βεβαρημένο οικογενειακό ιστορικό καρκίνου του μαστού και των ωθηκών εμφανίζουν μεταλλάξεις στα γονίδια BRCA1 και BRCA2, αν και το ποσοστό αυτό μπορεί να αλλάζει ανάλογα με το μελετούμενο πληθυσμό.

Όπως κι οι περισσότεροι τύποι καρκίνου, ο καρκίνος του μαστού είναι μια πολύπλοκη, πολυ-παραγοντική νόσος στην οποία υπάρχει μια έντονη αλληλεπίδραση μεταξύ γενετικών και περιβαλλοντικών παραγόντων. Είναι σημαντικό να αναφερθεί πως έχουν διαμορφωθεί συγκεκριμένες διεθνείς οδηγίες (American Society of Clinical Oncology (ASCO), National Comprehensive Cancer Network), που καθορίζουν αφενός τα κριτήρια γενετικού ελέγχου ειδικά για τα δυο αυτά γονίδια, αφετέρου τις συνιστώμενες ενέργειες ανάλογα με τα αποτελέσματα του ελέγχου. Όσο αφορά τον τρόπο κληρονομικότητας, στη συντριπτική πλειοψηφία των περιπτώσεων είναι αυτοσωματικός επικρατής, με μεταβλητή εκφραστικότητα και διεισδυτικότητα, που σημαίνει ότι άτομα με μετάλλαξη σε κάποιο από τα δυο αυτά γονίδια δεν θα εκδηλώσουν υποχρεωτικά τη νόσο ή θα την εκδηλώσουν με διαφορετικό τρόπο και βαρύτητα. Ειδικά για το γενετικό



έλεγχου των γονιδίων BRCA1 και BRCA2, δεν υπάρχουν συγκεκριμένες μεταλλάξεις, που να εμφανίζονται με ιδιαίτερα μεγάλη συχνότητα στους πάσχοντες, με εξαίρεση ίσως μεμονωμένες πληθυσμιακές ομάδες. Λαμβάνοντας υπόψη την αυξημένη σημασία του ελέγχου αλλά και την ετερογένεια της νόσου, ο μοριακός γενετικός έλεγχος μεταλλάξεων στα παραπάνω γονίδια, που είναι σχετικά μεγάλα σε μέγεθος, πρέπει να είναι όσο το δυνατό πληρέστερος, καλύπτοντας όλες τις πιθανές μεταλλάξεις με πλήρη ανάλυση της αλληλουχίας του DNA των γονιδίων αυτών. Ο στοχευμένος γενετικός έλεγχος των γονιδίων BRCA1 και BRCA2 αποτελεί το βασικό τεστ στον κληρονομικό καρκίνο του μαστού και των ωοθηκών. Ενώ, τα δύο αυτά γονίδια προσδίδουν έναν υπολογίσιμο κίνδυνο για καρκίνο του μαστού, πρόσφατες έρευνες έχουν εντοπίσει νέα γονίδια που επίσης συνδέονται με αυξημένο κίνδυνο εκδήλωσης καρκίνου του μαστού. Συγκεκριμένα, μια ομάδα από 15-20 γονίδια έχουν συσχετισθεί με επιπλέον ποσοστό έως 20% του κληρονομικού καρκίνου του μαστού, και ορισμένα από αυτά τα γονίδια συνδέονται επίσης με αυξημένο κίνδυνο για άλλους καρκίνους, όπως ο καρκίνος του παγκρέατος, καρκίνος των ωοθηκών και σαρκώματα. Τα γονίδια BARD1, BRIP1, NBN, RAD50 και RAD51C συμμετέχουν στον μηχανισμό επιδιόρθωσης του DNA αλληλεπιδρώντας με τα γονίδια BRCA1 και BRCA2, με συνέπεια μεταλλάξεις σε αυτά τα γονίδια να προσδίδουν μια αύξηση του κινδύνου για καρκίνο του μαστού κατά 4-5 φορές.

Επομένως, η αποκάλυψη μεταλλάξεων με το ταυτόχρονο γονιδιωματικό έλεγχο πολλαπλών γονιδίων (που συμπεριλαμβάνει και τα γονίδια BRCA1 και BRCA2), εκτιμάται ότι μπορεί να οδηγήσει σε 5-10 φορές αύξηση της εκτίμησης του κινδύνου για καρκίνο του μαστού.

#### **3.5.4 Περιγραφή των κλινικών εξετάσεων για την αντιμετώπιση του καρκίνου με χρήση της τεχνικής NGS**

Μία τυπική κλινική εξέταση με αλληλούχιση επόμενης γενιάς του γονιδιώματος του καρκινικού όγκου, έχει ένα τυποποιημένο λεπτομερές πρωτόκολλο το οποίο περιγράφει τις διαδικασίες πριν και μετά την ανάλυση. Οι διαδικασίες πριν την ανάλυση περιλαμβάνουν τα κλινικά χαρακτηριστικά του ασθενούς και λεπτομέρειες για τη συλλογή και την προετοιμασία των δειγμάτων (Marrone et al., 2014). Στη συνέχεια, οι μεταβλητές της ανάλυσης που θα χρησιμοποιηθούν μπορεί να επηρεάσουν την

ακρίβεια της αλληλούχισης του πάνελ των γονιδίων και σχετίζονται με την πλατφόρμα αλληλούχισης που θα χρησιμοποιηθεί. Οι μεταβλητές της μετα-ανάλυσης και της βιοπληροφορικής επεξεργασίας (εισαγωγή των δεδομένων, επικύρωση και ερμηνεία των αποτελεσμάτων) σχετικά με την αλληλούχιση επόμενης γενιάς περιλαμβάνουν μεταξύ άλλων κυρίως την κλινική ερμηνεία των αποτελεσμάτων και την αποθήκευση των δεδομένων σε παγκόσμιες βάσεις δεδομένων ελεύθερα προσβάσιμες στην Ιατρική κοινότητα (Forbes et al., 2011; Van Allen et al., 2013; Stratton et al., 2009; Wu et al., 2014). Η επεξεργασία των αποτελεσμάτων είναι σημαντική για την επιτυχή ολοκλήρωση της εφαρμογής και ορισμένες φορές ποικίλει ανάλογα και με την κλινική πρακτική, ενώ στηρίζεται κυρίως στην ευθυγράμμιση του γονιδιώματος από το καρκινικό δείγμα με το ανθρώπινο γονιδίωμα αναφοράς έτσι, ώστε να βρεθούν οι παραλλαγές (μεταλλάξεις, SNPs, αναδιοργανώσεις του γονιδιώματος) που σχετίζονται με το συγκεκριμένο είδος καρκίνου. Όλες οι βιοπληροφορικές αναλύσεις έχουν εξειδικευμένους υπολογιστικούς αλγόριθμους σχεδιασμένους να διαχειριστούν πολύ μεγάλους όγκους δεδομένων και αλληλουχιών, οι οποίες αρχικά χρησιμοποιούνταν μόνο για ερευνητικούς σκοπούς. Παρόλα αυτά, οι αναλύσεις αυτές έχουν πλέον αναβαθμισθεί με τέτοιο τρόπο ώστε να περιλαμβάνουν και γενικότερα δεδομένα για την κλινική αντιμετώπιση του καρκίνου (Van Allen et al., 2013).

Η διευρυμένη χρήση της αλληλούχισης επόμενης γενιάς για κλινικούς λόγους, παρέχει στους ερευνητές μια πολύ πιο ολοκληρωμένη εικόνα της γονιδιωματικής αρχιτεκτονικής των όγκων. Πλέον οι προσπάθειες να αξιοποιηθούν αυτές οι τεχνολογικές εξελίξεις στο έπακρον έχουν γίνει πραγματικότητα και έτσι διάφορες πλατφόρμες NGS παρέχουν εμπορικά τεστ ανάλυσης πάνελ γονιδίων για τον προσδιορισμό όγκων καρκίνου (Πίνακας 2).

Αν και υπάρχει μια έλλειψη ολοκληρωμένων δεδομένων για το ποιες μεταλλάξεις ή/και πολυμορφισμοί αποτελούν ευρήματα που χρήζουν κλινικής έρευνας έναντι του καρκίνου, έχουν δημιουργηθεί τρεις κατηγορίες πολυμορφισμών (Frueh et al., 2008):

- 1) Πολυμορφισμοί που συνδέονται με FDA-εγκεκριμένο φάρμακο για συγκεκριμένο τύπο όγκου,
- 2) Πολυμορφισμοί που συνδέονται με FDA-εγκεκριμένο φάρμακο χωρίς συσχέτιση με καρκινικό όγκο,
- 3) Πολυμορφισμοί που συνδέονται με μη εγκεκριμένα φάρμακα από τον FDA

και βρίσκονται σε πρόδρομες κλινικές έρευνες.

| Organization        | Tumor panel                | Markers                                     | Licenses and accreditation | Sample requirements                 | NGS platform                                | Analytic sensitivity | Sequence alignment                                              | Variant identification and molecular annotation | Turn-around time |
|---------------------|----------------------------|---------------------------------------------|----------------------------|-------------------------------------|---------------------------------------------|----------------------|-----------------------------------------------------------------|-------------------------------------------------|------------------|
| Arup Laboratories   | Solid Tumor Mutation Panel | 48 genes                                    | CLIA and CAP               | FFPE <10% tumor tissue not accepted | Not reported                                | Not reported         | Not reported                                                    | Not reported                                    | 12 to 14 days    |
| AsuraGen            | SuraSeq™                   | 17 genes and 500 annotated COSMIC mutations | CLIA and CAP               | FFPE or FNA                         | Ion Torrent and multiple Illumina platforms | 94% to 100%          | Proprietary sequencing data analysis performed using SuraScore™ | Proprietary variant caller: SuraSight™          | Not reported     |
| Foundation Medicine | FoundationOne™             | 236 genes; 47 introns from another 19 genes | CLIA and CAP               | FFPE ≥ 40 μm and ≥ 20% tumor tissue | Illumina HiSeq2000                          | 95% to 99%           | Not reported                                                    | Online Interactive Cancer Explorer™             | 14 to 17 days    |

CAP – College of American pathologists  
 CLIA – Clinical Laboratory Improvement Amendments  
 FFPE – formalin-fixed paraffin embedded  
 FNA – fine needle aspirates

**Πίνακας 3. Παραδείγματα εμπορικά διαθέσιμων πάνελ γονιδίων, για τον προσδιορισμό των καρκινικών τύπων. . (Marrone et al., 2014)**

### **3.5.4 Αλληλούχηση επόμενης γενιάς και ασθένειες που σχετίζονται με τη μιτοχονδριακή λειτουργία**

Οι ασθένειες που σχετίζονται με τη μιτοχονδριακή λειτουργία έχουν μεγάλη ποικιλομορφία και αυτό οφείλεται στο μοναδικό πρότυπο κληρονομικότητας τους λόγω των μιτοχονδρίων. Μπορούν να προέλθουν από μεταλλάξεις είτε στα πυρηνικά είτε στα μιτοχονδριακά γονίδια (Calvo et al., 2012). Η κλινική εικόνα αυτών των ασθενειών είναι οι πρωτογενείς σκελετικές και καρδιακές μυοπάθειες, οι νευρολογικές διαταραχές ή διαταραχές πολλών συστημάτων του οργανισμού ταυτόχρονα.

Υπάρχει ένα πάνελ 70 πυρηνικών γονιδίων το οποίο εμπλέκεται σε ασθένειες που σχετίζονται με τη μιτοχονδριακή λειτουργία και έτσι η τεχνική της αλληλούχησης επόμενης γενιάς μπορεί να ελέγξει και τα 70 γονίδια ταυτόχρονα ανιχνεύοντας πιθανές μεταλλάξεις και εξάγοντας σημαντικά συμπεράσματα για τους ασθενείς (Calvo et al., 2012). Επιπρόσθετα, σε μια άλλη μελέτη ελέγχθηκε ένα πάνελ γονιδίων που περιείχε όλα τα μιτοχονδριακά και περίπου άλλα 1,000 πυρηνικά γονίδια από 42 ανεξάρτητα

άτομα μικρής ηλικίας που παρουσίαζαν κλινικά και βιοχημικά ευρήματα ασθένειας που σχετίζεται με την μιτοχονδριακή οξειδωτική φωσφορυλίωση (Vasta et al., 2009). Το 55% των ασθενών είχε μεταλλάξεις κυρίως στο μιτοχονδριακό γονιδίωμα και σε ορισμένα πυρηνικά γονίδια, που σχετίζονται με την παραπάνω ασθένεια.

### **3.5.5 Αλληλούχιση επόμενης γενιάς και καρδιαγγειακές παθήσεις**

Οι καρδιακές αρρυθμίες είναι παγκοσμίως μια από τις κύριες αιτίες νοσηρότητας και θνησιμότητας. Οι συγγενείς καρδιακές αρρυθμίες αποτελούν μια ξεχωριστή ομάδα από καρδιακές διαταραχές, που προκύπτουν από αλλοιώσεις στις ηλεκτροφυσιολογικές ιδιότητες της καρδιάς. Συγκεκριμένα, η συντονισμένη καρδιακή δραστηριότητα περιλαμβάνει, μεταξύ άλλων, το συγχρονισμένο και διαδοχικό άνοιγμα και κλείσιμο διαύλων ιόντων σε απόκριση προς το ηλεκτρικό δυναμικό και μεταδίδει το δυναμικό δράσης σε κάθε διαμέρισμα της καρδιάς.

Παρά το γεγονός ότι οι περιβαλλοντικοί παράγοντες συμβάλλουν σαφώς στην αρρυθμιογένεση, οικογενειακές και πληθυσμιακές μελέτες έχουν αποδείξει την ύπαρξη γενετικής αιτιολογίας. Για παράδειγμα, μεταλλάξεις σε >20 γονίδια, που κωδικοποιούν και ρυθμίζουν ειδικά κανάλια ιόντων, συνδέονται με διάφορες μορφές αρρυθμιολογικών διαταραχών, που συμβαίνουν σε μια κατά τα άλλα δομικά φυσιολογική καρδιά (Meder et al., 2011). Μια τέτοια πάθηση είναι για παράδειγμα και το σύνδρομο Brugada (BRS), που συνδέεται με μεταλλάξεις σε τουλάχιστον 8 διαφορετικά γονίδια και χαρακτηρίζεται από αυξημένο κίνδυνο θανατηφόρων κοιλιακών αρρυθμιών. Τα γονίδια αυτά κωδικοποιούν πρωτεΐνες-κανάλια καλίου και νατρίου και είναι τα *CN5A*, *GPD1L*, *CACNA1C*, *CACNB2*, *SCN1B*, *KCNE3*, *SCN3B*, και *HCN4*. Ο αιφνίδιος καρδιακός θάνατος (Sudden Cardiac Death - SCD) (αιφνίδιος θάνατος άγνωστης αιτιολογίας), είναι συχνή αιτία θνησιμότητας, που επηρεάζει όλες τις ηλικίες. Η βασική αιτία αιφνίδιου καρδιακού θανάτου σε άτομα μεγαλύτερα των 45 ετών οφείλεται κυρίως στην αθηροσκληρωτική στεφανιαία νόσο.

Γενικά, η πλειοψηφία των καρδιαγγειακών νοσημάτων εκδηλώνονται και κληρονομούνται με τον αυτοσωματικό επικρατή τρόπο, που σημαίνει ότι άτομα με μετάλλαξη σε ένα μόνο από τα δυο αντίγραφα κάποιου γονιδίου έχουν ιδιαίτερα αυξημένο κίνδυνο να νοσήσουν κι όλοι οι πρώτου βαθμού συγγενείς ενός νεαρού ασθενούς έχουν 50% κίνδυνο να κληρονομήσουν την ασθένεια (Meder et al., 2011).

Επίσης, είναι σημαντικό να αναφερθεί ότι διαφορετικές μεταλλάξεις στο ίδιο γονίδιο μπορεί να οδηγήσουν στην εκδήλωση διαφορετικού τύπου καρδιαγγειακού νοσήματος.

Οι μέχρι σήμερα διαθέσιμες επιλογές γενετικού ελέγχου καρδιαγγειακών νοσημάτων ήταν εξαιρετικά αργές και ελλιπείς (έλεγχος μόνο για λίγα γονίδια), με αποτέλεσμα οι γενετικές αιτίες - μεταλλάξεις συχνά να μην αναγνωρίζονται σε ασθενείς με συγγενείς και κληρονομικές παθήσεις της καρδιάς.

Στο βαθμό λοιπόν, που ο μέχρι σήμερα γενετικός έλεγχος ήταν επιλεκτικός, ατελής και συχνά δαπανηρός, ο μαζικός έλεγχος όλων των γνωστών γονιδίων με γονιδιωματική ανάλυση οδηγεί με ένα μόνο βήμα στην επιτυχή κι οριστική διάγνωση. Η αποκάλυψη της παθολογικής μετάλλαξης έχει μεγάλη σημασία για την γενετική συμβουλευτική των ασθενών και των οικογενειών, διευκολύνοντας τη διάγνωση σε άτομα που διατρέχουν κίνδυνο. Παράλληλα, σε πολλές περιπτώσεις η ακριβής γνώση της γενετικής βλάβης οδηγεί στην αποτελεσματικότερη διαχείριση των συμπτωμάτων καθώς και στον προσδιορισμό της δέουσας εξατομικευμένης φαρμακευτικής αγωγής ή θεραπείας.

### **3.5.6 Αλληλούχιση επόμενης γενιάς και αμφιβληστροειδοπάθειες**

Γενετικές ασθένειες του οφθαλμού ή και με τη συμμετοχή του οφθαλμού είναι σχετικά κοινές, αποτελώντας την κύρια αιτία τύφλωσης σε παιδιά και ενήλικες, κι εμφανίζουν μεγάλη κλινική και γενετική ετερογένεια. Ένας αριθμός από ασθένειες του αμφιβληστροειδούς, όπως η μελαγχρωστική αμφιβληστροειδοπάθεια, η κωνίων-ραβδίων δυστροφία, η αχρωματοψία και άλλες προκαλούν εκφυλισμό του φωτοϋποδοχέα και συχνά είναι δύσκολο να διακριθεί ο συγκεκριμένος τύπος. Έχουν περιγραφεί πάνω από 90 μορφές μελαγχρωστικής αμφιβληστροειδοπάθειας κι έχουν εντοπισθεί περισσότερα από 60 υπεύθυνα γονίδια (Daiger et al., 2010). Αντίστοιχα, η συγγενής αμαύρωση Leber (LCA) οφείλεται σε περίπου 12-15 γονίδια, χαρακτηρίζεται από απώλεια όρασης από τη γέννηση και με την πάροδο του χρόνου μπορεί να εκδηλωθεί κολόβωμα της ωχράς κηλίδας και αμφιβληστροειδοπάθεια. Οι περισσότερες οφθαλμικές γενετικές διαταραχές δεν είναι ακόμη αντιμετωπίσιμες και δεν έχουν θεραπευτικές προσεγγίσεις, κυρίως λόγω της περιορισμένης κατανόησης της παθογένεσης.

Με δεδομένη τη φύση, τη σοβαρότητα αλλά και την πολυπλοκότητά τους, ο γονοδιωματικός έλεγχος με αλληλούχιση επόμενης γενιάς για τα οφθαλμολογικά γενετικά νοσήματα είναι ιδιαίτερα πολύτιμος, αν όχι απαραίτητος, κι έχει ως στόχο την ακριβή και αποτελεσματική ταυτοποίηση των γενετικών αιτίων. Η αναγνώριση του συγκεκριμένου οφθαλμολογικού νοσήματος μέσω του γενετικού ελέγχου, αναμφίβολα διευκολύνει την ανάπτυξη και την εισαγωγή νέων και εξατομικευμένων θεραπειών.

Επίσης έχει αναφερθεί πως μεταλλάξεις σε περισσότερα από 150 γονίδια προκαλούν κληρονομήσιμες διαταραχές του αμφιβληστροειδούς, που οδηγούν σε μια ποικιλομορφία αλληλοεπικαλυπτόμενων φαινοτύπων (Audo et al., 2012). Ενώ η εξέταση κάθε επιμέρους γονιδίου είναι χρονοβόρα και οικονομικά απαγορευτική, οι Audo et al. (2012) εφήρμοσαν μέσω αλληλούχισης επόμενης γενιάς την ανάλυση μια περιοχής 1.177 Mb, που περιελάμβανε 254 γονίδια που σχετίζονται άμεσα με τη λειτουργία του αμφιβληστροειδούς. Οι ασθενείς που συμμετείχαν στην έρευνα εμφάνιζαν οφθαλμολογικά γενετικά νοσήματα όπως μελαγχρωστική αμφιβληστροειδοπάθεια, κωνίων-ραβδίων δυστροφία, αχρωματοψία, δυστροφία της ωχράς του Stargardt, κ.α. Χαρακτηρίστηκαν τρεις γνωστές και πέντε νέες μεταλλάξεις των γονιδίων *NR2E3*, *PRPF3*, *EYS*, *PRPF8*, *CRB1*, *TRPM1* και *CACNA1F*. Η παραπάνω προσέγγιση είναι πολύ αποδοτική χρονικά και έδωσε νέα αποτελέσματα συσχέτισης των φαινοτύπων λόγω των νόσων με τα συγκεκριμένα γονίδια προς ανάλυση (Audo et al., 2012). Μια άλλη προσέγγιση είναι αυτή των Daiger et al., (2010) – (VisionChip), όπου μικροσυστοιχίες με ολιγονουκλεοτίδια δεσμεύουν κατακερματισμένο DNA και στη συνέχεια πραγματοποιείται αλληλούχιση επόμενης γενιάς για το έλεγχο 593 γονιδίων. Αυτή η μέθοδος είναι πιθανό να μειώσει αρκετά ακόμα το κόστος ανά ασθενή με διαταραχές του αμφιβληστροειδούς.

### **3.5.7 Αλληλούχιση επόμενης γενιάς και αιματολογικές διαταραχές**

Η διάγνωση μορφών της πρωτογενούς ανοσολογικής ανεπάρκειας είναι δύσκολη, κυρίως λόγω της παρόμοιας κλινικής εικόνας για κάθε μορφή. Η πρωτογενής ανοσολογική ανεπάρκεια περιλαμβάνει μια πλειάδα σπάνιων και συχνά κληρονομικών παθήσεων (π.χ. σύνδρομο DiGeorge, σύνδρομο Omenn, ανεπάρκεια ανοσοφαιρίνης A, ανεπάρκεια του παράγοντα 3 του συστήματος του συμπληρώματος.). Οι ασθενείς παρουσιάζουν μειωμένη ικανότητα αντιμετώπισης παθογόνων μικροοργανισμών από

την παιδική ηλικία. Η σοβαρότητα της νόσου εξαρτάται από το ποιο στοιχείο του ανοσοποιητικού συστήματος απουσιάζει ή υπολειτουργεί. Η αντιμετώπιση είναι αυστηρώς εξειδικευμένη και περιλαμβάνει χορήγηση ανοσοσφαιρινών, αντιβιοτικών, μεταμόσχευση μυελού των οστών και την αναπτυσσόμενη γονιδιακή θεραπεία. Οι κλασικές τεχνικές διάγνωσης περιλαμβάνουν κυτταρολογικές αναλύσεις οι οποίες δεν είναι πάντα αποδοτικές. Οι Ghosh et al. (2012) χρησιμοποιώντας την πλατφόρμα GS FLX Titanium 454, αλληλούχισαν 395 γονίδια που σχετίζονται με διάφορες μορφές πρωτογενούς ανοσολογικής ανεπάρκειας και χαρακτήρισαν συγκεκριμένες μεταλλάξεις. Με αυτό τον τρόπο δημιούργησαν ένα σημαντικό εργαλείο-βάση δεδομένων ως προς τη διάγνωση της πρωτογενούς ανοσολογικής ανεπάρκειας.

Παρόμοια με την πρωτογενή ανοσολογική ανεπάρκεια, οι διαταραχές των αιμοπεταλίων είναι γενοτυπικά διαφορετικές εμφανίζοντας σχεδόν ίδια κλινικά συμπτώματα. Προκειμένου να γίνει επιτυχής διάγνωση του είδους της διαταραχής έχει αναφερθεί η εφαρμογή αλληλούχισης επόμενης γενιάς με την πλατφόρμα Illumina για την ανάλυση 216 πιθανών γονιδίων (Jones et al., 2012). Η λίστα αυτή περιλαμβάνει τα γονίδια *HPS1*, *HPS1*, *VPS33B*, *NAPA*, *LYST*, *HPS4*, *VPS18*, *VPS16*, *SCL3N* και ενώ βρέθηκαν 4.500 διαφορετικοί πολυμορφισμοί σε αυτά τα 216 γονίδια, η μετα-ανάλυση ανέδειξε 10 πολυμορφισμούς οι οποίοι σχετίζονται άμεσα με την πάθηση (Jones et al., 2012). Η παραπάνω μελέτη αποδεικνύει τη χρησιμότητα της τεχνικής NGS στη ταχεία διάγνωση της διαταραχής των αιμοπεταλίων με τη χρήση ενός μόνο τεστ.

### **3.5.8 Αλληλούχιση επόμενης γενιάς και πρωτοπαθής δυσκινησία των κροσσών**

Η πρωτοπαθής δυσκινησία των κροσσών είναι πολυσυστηματική νόσος εκφραζόμενη κλινικά από αντίστοιχη παθολογία των συστημάτων και των επί μέρους οργάνων που διαθέτουν ανάλογο επιθήλιο. Συγκεκριμένα, από το αναπνευστικό και τις παραρρινίες κοιλότητες παρατηρούνται υποτροπιάζουσες αναπνευστικές λοιμώξεις, βρογχεκτασίες και επεισόδια ωτίτιδας και ιγμορείτιδας. Από το γεννητικό σύστημα αμφοτέρων των φύλων παρατηρείται υπογονιμότητα. Το σύνδρομο συνδέεται σε 50% περίπου των ασθενών με αναστροφή των σπλάγγων (σύνδρομο Kartagener). Η διάγνωση γίνεται με μια σειρά μη ειδικών εξετάσεων όπως δοκιμασία σακχαρίνης, μέτρησης της συχνότητας κίνησης των κροσσών και απεικόνιση της μορφολογίας των

κροσσών και του προσανατολισμού των με το ηλεκτρονικό μικροσκόπιο. Επιπρόσθετα, η ανοσοιστοχημεία μπορεί να βοηθήσει στη διάγνωση μέσω της ανίχνευσης συγκεκριμένων πρωτεϊνών (παρουσία ή έλλειψή τους), αλλά και με αυτό τον τρόπο δεν μπορεί να βρεθεί μια γενετική αιτιολογία. Έχει αναφερθεί η χρήση της τεχνικής NGS (GS FLX Titanium Roche 454 πλατφόρμα) στη μελέτη ασθενών με την παραπάνω νόσο και έγινε αλληλούχιση 2089 εξωνίων από 79 πιθανά συσχετιζόμενα γονίδια (Berg et al., 2011). Βρέθηκαν και χαρακτηρίστηκαν πολλαπλές μεταλλάξεις στα γονίδια *DNAH5*, *DNAH11*, *DNAI1*, *DNAI2*, *KTU*, *LRR50*, *RSPH9*, *RSPH4A* και *TXNDC3*, αλλά και σημαντικές γονιδιακές αναδιατάξεις. Σε μία άλλη παρόμοια μελέτη πραγματοποιήθηκε ανάλυση ενός πάνελ 70 γονιδίων και βρέθηκε ότι συγκεκριμένες μεταλλάξεις στα γονίδια *RSPH1*, *RSPH9*, *RSPH4A* ευθύνονται για την πρωτοπαθή δυσκινησία των κροσσών (Onoufriadis et al., 2014a). Επίσης, η εφαρμογή της τεχνικής NGS για αλληλούχιση ολόκληρου του γονιδιώματος στη μελέτη δύο οικογενειών με ιστορικό στην πρωτοπαθή δυσκινησία των κροσσών, αποκάλυψε μη-νοηματικές μεταλλάξεις (non-sense mutation) στο γονίδιο *ARMC4*, προσθέτοντάς το στο πάνελ των γονιδίων που σχετίζονται με τη νόσο (Onoufriadis et al., 2014b).

### **3.5.9 Αλληλούχιση επόμενης γενιάς και επιληψία**

Ως επιληψία ορίζεται μία ομάδα νευρολογικών διαταραχών που χαρακτηρίζονται από παροξυσμική, αυτοδιατηρούμενη, ακανόνιστα επαναλαμβανόμενη, υπέρμετρη και υπερσύγχρονη ενεργοποίηση σε μία μόνο περιοχή (εστιακή επιληψία) ή σε πολλές περιοχές ταυτόχρονα (γενικευμένη επιληψία) (Privitera, 2011). Η επιληψία συνήθως ταξινομείται με βάση τις κλινικές και ηλεκτροεγκεφαλογραφικές εκδηλώσεις των κρίσεων, οι οποίες ορίζονται ως παροξυσμική, παροδική δυσλειτουργία του κεντρικού νευρικού συστήματος (Κ.Ν.Σ.). Οι επιληπτικές κρίσεις φαίνεται να ξεκινούν συνήθως σε περιοχές του φλοιού και διακρίνονται στις *εστιακές* (partial), οι οποίες ξεκινούν σε συγκεκριμένο σημείο του φλοιού (εστία) και στις *γενικευμένες* (generalized). Οι εκδηλώσεις συμπεριφοράς των κρίσεων εξαρτώνται από τη λειτουργία η οποία φυσιολογικά ελέγχεται από την περιοχή εκείνη του φλοιού που η κρίση συμβαίνει (McNamara, 1994; McCormick and Contreras, 2001). Η επιληψία είναι ένα πολυπαραγοντικό σύνδρομο όπου μια τεράστια ποικιλία από περιβαλλοντικά αλλά και



γενετικά αίτια συντελούν στην εμφάνιση της (McCormick and Contreras, 2001). Ο συνδυασμός των παραπάνω αιτίων οδηγεί στην πληθώρα των ειδών των επιληπτικών κρίσεων και της διαφορετικής συμπεριφορικής έκφρασης τους τόσο μεταξύ των ατόμων, όσο και στο ίδιο άτομο.

Οι επιληπτικές κρίσεις μπορεί να έχουν περιορισμένη διάρκεια ή όχι και τυπικά ελέγχονται με την κατάλληλη θεραπευτική αγωγή. Σε ορισμένες περιπτώσεις, ωστόσο, άτομα με επιληπτική εγκεφαλοπάθεια έχουν σοβαρές επιληπτικές κρίσεις, που δεν ανταποκρίνονται σε θεραπεία και οδηγούν σε σοβαρές διαταραχές. Ο γονιδιακός έλεγχος στις επιληψίες έχει ως στόχο την ακριβή και αποτελεσματική ταυτοποίηση των γενετικών αιτίων και τα αποτελέσματα είναι ιδιαίτερα χρήσιμα για τη διάγνωση του συγκεκριμένου τύπου επιληψίας, για την πρόβλεψη της εξέλιξης της νόσου, για τον προσδιορισμό της ενδεδειγμένης φαρμακευτικής αγωγής και τη διαχείριση των συμπτωμάτων και για πρόληψη της νόσου στην ευρύτερη οικογένεια (Wang et al., 2014). Η αναγνώριση του συγκεκριμένου μηχανισμού εκδήλωσης της επιληψίας μέσω του γενετικού ελέγχου, διευκολύνει αναμφίβολα την ανάπτυξη και την εισαγωγή νέων και εξατομικευμένων θεραπειών. Η εξέταση με χρήση NGS περιλαμβάνει την ανάλυση 300 γονιδίων που έχουν συσχετισθεί με όλες τις γνωστές μορφές επιληψίας (Wang et al., 2014).

### **3.5.10 Αλληλούχιση επόμενης γενιάς και διαταραχές της ακοής**

Οι διαταραχές της ακοής διακρίνονται σε επίκτητες (λόγω περιβαλλοντικών παραγόντων, αλλά και σε λοιμώξεις του εμβρύου π.χ. τοξόπλασμα, ερυθρά, κυτταρομεγαλοϊός), σε αγνώστου αιτιολογίας και σε γενετικές. Η γενετική διαταραχή της ακοής δύναται να είναι μέρος ενός συνδρόμου ή όχι, ακολουθώντας και τους τρεις πιθανούς τύπους κληρονομικότητας, την υπολειπόμενη, επικρατή και φυλοσύνδετη. Η μη-συνδρομική διαταραχή ακοής και ειδικότερα η μορφή που εκδηλώνεται πριν την ανάπτυξη της γλωσσικής επικοινωνίας στα παιδιά, αποτελεί σε μεγάλο ποσοστό γενετικό νόσημα, που οφείλεται κύρια σε μεταλλάξεις υπολειπόμενου χαρακτήρα με μεγάλη συχνότητα φορέων στον πληθυσμό. Είναι πολύ σημαντικό για τη σωστή διαχείριση των ασθενών, αλλά και για την γενετική καθοδήγηση σε μελλοντικές κυήσεις, να μπορούμε να διαχωρίσουμε τις κληρονομικές μορφές των διαταραχών από τις

επίκτητες (μη-γενετικές) μορφές. Αυτό επιτυγχάνεται με συνδυασμό διάφορων κλινικών εξετάσεων, καθώς και με αντίστοιχες γενετικές εξετάσεις.

Περισσότερα από 100 γονίδια υπολογίζεται ότι ενέχονται στις διαταραχές της ακοής. Ωστόσο, μεταλλάξεις σε δύο γονίδια, της κοννεξίνης 26 (*GJB2*) και της κοννεξίνης 30 (*GJB6*), αντιπροσωπεύουν ποσοστό περισσότερο από 65% της μη-συνδρομικής βαρηκοΐας και περίπου 35% της σποραδικής μορφής. Μάλιστα, συγκεκριμένα για το γονίδιο της κοννεξίνης 26 (*Cx26*), μια μετάλλαξη που ονομάζεται 35delG ανιχνεύεται σε συχνότητα ~90% των παθολογικών μεταλλάξεων του γονιδίου και με συχνότητα φορέων στο γενικό πληθυσμό περίπου 3,5%.

Ο μοριακός γενετικός έλεγχος των γονιδίων *GJB2* και *GJB6* είναι ο πρώτος και βασικός έλεγχος που πρέπει να εκτελείται σε άτομα με μη-συνδρομική διαταραχή της ακοής. Η διάγνωση βασίζεται στον έλεγχο μεταλλάξεων όλου του γονιδίου *GJB2*, καθώς και στον πλήρη έλεγχο ελλείψεων/αναδιπλασιασμών του γονιδίου *GJB6*, με αποτέλεσμα να καλύπτεται ποσοστό >99% των μεταλλάξεων της αυτοσωματικής υπολειπόμενης μη-συνδρομικής βαρηκοΐας.

Σε περίπτωση κατά την οποία ο παραπάνω έλεγχος δεν αποκαλύψει τη γενετική αιτία της κώφωσης είναι εφικτή η εφαρμογή μαζικού ελέγχου μεταλλάξεων σε περισσότερα από 80 γονίδια, που έχουν συσχετισθεί με όλες τις μορφές διαταραχών της ακοής, ελέγχοντας γονίδια που σχετίζονται με συνδρομικές, αλλά και μη-συνδρομικές μορφές της νόσου (Sivakumaran et al., 2013). Οι στρατηγικές που εφαρμόζονται για το γενετικό έλεγχο έχουν ως στόχο την ακριβή και αποτελεσματική ταυτοποίηση των γενετικών αιτίων και τα αποτελέσματα χρησιμοποιούνται για περαιτέρω επεξήγηση της αιτίας, για πρόβλεψη των ακουστικών χαρακτηριστικών, για πρόληψη της νόσου στην οικογένεια, για τη διαχείριση των συμπτωμάτων, καθώς και για τον προσδιορισμό της θεραπείας. Η αναγνώριση του συγκεκριμένου μηχανισμού αλλοίωσης των κυττάρων στο εσωτερικό του αυτιού μέσω του γενετικού ελέγχου, αναμφίβολα διευκολύνει την ανάπτυξη και την εισαγωγή νέων και εξειδικευμένων θεραπειών για διαφορετικές μορφές διαταραχής της ακοής (Sivakumaran et al., 2013). Λόγω αυτών των ιδιαίτερων χαρακτηριστικών της νόσου, ο γενετικός έλεγχος έχει εξαιρετικά υψηλή αξία ως κλινική εφαρμογή και αποτελεί σημαντικό εργαλείο στη διάγνωση.

## 3.6 Αλληλούχηση επόμενης γενιάς και επιγενετική

### 3.6.1.Γενικά στοιχεία

Ο όρος επιγενετική χρησιμοποιήθηκε αρχικά για να περιγράψει φαινόμενα όπου οι αλληλεπιδράσεις μεταξύ γενοτύπου και περιβάλλοντος μπορούν να παράγουν νέους φαινοτύπους κατά τη διάρκεια της ανάπτυξης (Waddington, 1953). Σήμερα, ο όρος επιγενετική αναφέρεται πιο ειδικά στη μελέτη κληρονομήσιμων τροποποιήσεων στην έκφραση των γονιδίων ή του φαινοτύπου των κυττάρων χωρίς αλλαγές στη νουκλεοτιδική αλληλουχία του DNA, που μπορούν να τροποποιηθούν από περιβαλλοντικά ερεθίσματα (Holliday, 1994). Η επιγενετική ρύθμιση της γονιδιακής έκφρασης επιτυγχάνεται με μεθυλίωση του DNA, τροποποιήσεις ή παραλλαγές των ιστονών της χρωματίνης, αναδιαμόρφωση της χρωματίνης και μπορεί επίσης να περιλαμβάνει και μικρά RNAs. Η μεθυλίωση του DNA που λαμβάνει χώρα στα νουκλεοτίδια των κυτοσινών, πραγματοποιείται από ένζυμα που ονομάζονται DNA μεθυλοτρανσφεράσες και εμπλέκεται σε πολλές κυτταρικές διαδικασίες όπως είναι η αποσιώπηση μεταθετών στοιχείων και περικεντρομερικών περιοχών, η απενεργοποίηση του X χρωμοσώματος και το γονιδιακό εντύπωμα. Οι τροποποιήσεις των ιστονών αναφέρονται στη μετα-μεταφραστική ομοιοπολική πρόσδεση χημικών ομάδων στις ιστόνες, όπως φωσφορυλίωση, ακετυλίωση και μεθυλίωση. Οι παραλλαγές των ιστονών είναι μη κανονικές ιστόνες με διαφορετικές αμινοξικές ακολουθίες από τις κανονικές ιστόνες και μπορεί να έχουν διαφορετικές επιγενετικές επιδράσεις στο γονιδίωμα από τις κανονικές ιστόνες. Σταθερές χρωματινικές δομές, που διατηρούνται ή τροποποιούνται από πρωτεΐνες αναδιαμόρφωσης της χρωματίνης, μπορεί επίσης να διαδραματίζουν σημαντικό ρόλο στη ρύθμιση της γονιδιακής έκφρασης. Τα μικρά RNAs διαδραματίζουν ποικίλους ρόλους στη ρύθμιση της γονιδιακής έκφρασης σε προ- αλλά και μετα-μεταφραστικό επίπεδο. Γενικά υποστηρίζεται πως οι επιγενετικοί μηχανισμοί διαδραματίζουν ουσιαστικό ρόλο σε όλους τους πολύπλοκους οργανισμούς ως ρυθμιστές της μεταγραφής.

Η μελέτη της λειτουργίας της μεθυλίωσης έδειξε ότι η μεθυλίωση του DNA συνδέεται με καταστολή της χρωματίνης και απαιτείται για τη σταθερότητα του γονιδιώματος (σίγαση των μεταθετών στοιχείων), τη ρύθμιση της γονιδιακής έκφρασης και τη διατήρηση της ακεραιότητας της ταυτότητας των κυττάρων (Suzuki και Bird, 2008). Φαίνεται να περιορίζεται σε επίπεδο κυττάρου και ειδικά στο πλαίσιο των CpG

δινουκλεοτιδίων. Τα CpG δινουκλεοτίδια είναι ακολουθίες κυτοσίνης-φωσφόρου-γουανίνης που συνήθως συσσωρεύονται σε γονιδιωματικές περιοχές που ονομάζονται CpG νησίδες και συχνά τοποθετούνται στους προαγωγείς των γονιδίων, εκδηλώνοντας χαμηλά επίπεδα DNA μεθυλίωσης. Στον άνθρωπο, το 60% περίπου των γονιδίων περιέχουν CpG νησίδες στην περιοχή του προαγωγέα τους και αν και η μεθυλίωση των περιοχών αυτών είναι σπάνια, όταν συμβαίνει έχει ως αποτέλεσμα τη σταθερή καταστολή του γονιδίου (Weber and Schübeler, 2007).

Η κληρονομικότητα των επιγενετικών χαρακτηριστικών ενός οργανισμού, που επηρεάζουν το φαινοτυπικό αποτέλεσμα, εξαρτάται από πληθώρα παραγόντων. Τα πρότυπα της μεθυλίωσης του DNA μπορεί να επηρεαστούν από τη γενετική παραλλακτικότητα, τις περιβαλλοντικές αλλαγές και από κληρονομήσιμες και μη αλλαγές σε άλλες επιγενετικές διαδικασίες όπως είναι, για παράδειγμα, η επίδραση που μπορεί να ασκήσει η δομή της χρωματίνης ή η πρόσδεση μεταγραφικών παραγόντων στα πρότυπα μεθυλίωσης του DNA. Επίσης, στοχαστικές τροποποιήσεις κατά το πέρασμα του χρόνου μπορούν να επηρεάσουν την επιγενετική κατάσταση του οργανισμού. Όλοι αυτοί οι παράγοντες μπορεί να συνεισφέρουν στην εκτίμηση της κληρονομικότητας της μεθυλίωσης του DNA, η οποία μπορεί να είναι χρόνο-, ιστο-, τόπο- ή πληθυσμό-ειδική (Bell και Spector, 2011).

Ύστερα από τον αρχικό ενθουσιασμό της αλληλούχισης ολόκληρου του γονιδιώματος, έγινε κατανοητό ότι αυτό από μόνο του και χωρίς επιπρόσθετες πληροφορίες δεν θα αποτελούσε το ιδανικό εργαλείο στα χέρια των επιστημόνων. Το επιγένωμα είναι διαφορετικό όχι μόνο από άτομο σε άτομο, αλλά και μέσα στον ίδιο τον οργανισμό ανάλογα με το αναπτυξιακό στάδιο και τον μελετώμενο ιστό. Έχει πραγματοποιηθεί μια συλλογική προσπάθεια αποκρυπτογράφησης του επιγνώματος και έχει αποτυπωθεί στο πρόγραμμα ENCODE (Encyclopedia of DNA elements). Αν και κάθε άτομο έχει ένα ξεχωριστό γονιδίωμα, σε αυτό υπάρχουν εκατοντάδες διαφορετικά επιγνώματα στα διάφορα αναπτυξιακά στάδια και ιστούς. Οι πρωτοφανείς εξελίξεις στην αλληλούχιση επόμενης γενιάς, λόγω του χαμηλού κόστους και του ελάχιστου χρόνου πραγματοποίησης (που συνεχίζουν να βελτιώνονται) έχουν οδηγήσει τη μοριακή βιολογία και τη γενετική έρευνα, σε καινοτόμους τρόπους ανάλυσης των κυτταρικών γεγονότων που αφορούν το DNA ή το RNA σε ένα επίπεδο επιγνώματος (Sarda and Hannenhalli, 2014). Οι κλινικές έρευνες που αφορούν τη χρήση της αλληλούχισης επόμενης γενιάς σε σχέση με το επιγένωμα και τις ασθένειες είναι εν

εξελιξεί και στο μέλλον θα βρίσκουν κυρίως εφαρμογή στις καρκινικές παθήσεις, όπου το επιγένομα παίζει σημαντικό ρόλο (Sarda and Hannenhalli, 2014).

## **3.7 Πρωτογενής ανάλυση δεδομένων αλληλούχισης επόμενης γενιάς**

### **3.7.1.Γενικά στοιχεία**

Όπως με όλες τις νεωτεριστικές τεχνολογίες στις επιστήμες της ζωής, το μέγεθος και πολλές φορές και η πολυπλοκότητα των παραγομένων συνόλων δεδομένων αυξάνεται εκθετικά όσο εξελίσσεται και γίνεται πιο προσβάσιμη η αντίστοιχη τεχνολογία, απαιτώντας πιο αποδοτικούς τρόπους για αποτελεσματική εκμετάλλευση των δεδομένων, σχηματίζοντας έτσι νέες υπολογιστικές και στατιστικές προκλήσεις στο πεδίο της βιοπληροφορικής. Στις περισσότερες περιπτώσεις, το παρεχόμενο λογισμικό από τους κατασκευαστές πλατφορμών NGS πραγματοποιεί μόνο τα θεμελιώδη βήματα βιοπληροφορικής ανάλυσης, καλύπτοντας διάφορες όψεις, όπως η ανάλυση εικόνας ανά κύκλο αλληλούχισης για την αναγνώριση των εντάσεων των αλληλουχηθέντων ομάδων. Οι εξαχθείσες εντάσεις χρησιμοποιούνται έπειτα για την αναγνώριση νουκλεοτιδίων σε συγκεκριμένες θέσεις, οδηγώντας έτσι στη δημιουργία μικρών αναγνωσθέντων θραυσμάτων αλληλουχιών (που θα αναφέρονται από εδώ και πέρα ως μικρές αλληλουχίες) τα οποία χαρτογραφούνται σε κάποιο γονιδίωμα αναφοράς. Ωστόσο, πηγές πολυπλοκότητας όπως οι διαφορές στη χημεία και στη συλλογή των ακατέργαστων δεδομένων μεταξύ των πλατφορμών NGS απαιτούν εξειδικευμένες γραμμές υπολογιστικής ανάλυσης και εμποδίζουν το συνδυασμό των αποτελεσμάτων (McPherson, 2009). Οι πλατφόρμες Illumina Genome Analyzer και Helicos HeliScope δημιουργούν εντάσεις σήματος που είναι συγκεκριμένες ανά νουκλεοτίδιο και βασικοί αλγόριθμοι προσδιορίζουν το πιο πιθανό επόμενο νουκλεοτίδιο στην αλληλουχία βάσει συγκεκριμένου πλαισίου, χτίζοντας έτσι μια αλληλουχία από νουκλεοτίδια σαν έξοδο. Σε αντίθεση, το σύστημα της Roche προσθέτει στη σχηματιζόμενη αλληλουχία μόνο έναν τύπο νουκλεοτιδίου κάθε φορά σχηματίζοντας σειρές μονονουκλεοτιδίων δημιουργώντας σήματα ανάλογα του αριθμού των νουκλεοτιδίων που μπορούν να μετατραπούν σε αλληλουχίες αλλά με κάποια αβεβαιότητα. Τέτοιες βασικές διαφορές

στη συλλογή δεδομένων σε συνδυασμό με το μεγάλο όγκο των δεδομένων που παράγονται, έχουν οδηγήσει σε ένα κενό μεταξύ παραδοσιακών εργαλείων ανίχνευσης/ανάλυσης αλληλουχιών και των δεδομένων NGS (McPherson, 2009). Καθώς κάθε πλατφόρμα έχει ανόμοιες εξόδους και μοναδικά προφίλ πιθανών σφαλμάτων, η εγκαθίδρυση μιας καθολικής ροής εργασιών για την εύρεση των νουκλεοτιδίων (κλήση βάσεων) και την ανάλυση των αλληλουχιών παραμένει μια πρόκληση.

Για να περιπλακεί περισσότερο το τοπίο της αλληλούχισης επόμενης γενιάς, αυτές οι πλατφόρμες εφαρμόστηκαν γρήγορα σε πολλές μεθόδους γονιδιωματικής οι οποίες παραδοσιακά δε χρησιμοποιούσαν απευθείας δεδομένα αλληλουχιών. Χαρακτηριστικό παράδειγμα είναι η αντικατάσταση της παραδοσιακής παρακολούθησης του μεταγραφώματος από μικροσυστοιχίες επειδή ο μεγάλος αριθμός αλληλουχιών που παράγονται σε έναν- και μόνο κύκλο λειτουργίας μιας πλατφόρμας NGS ανταγωνίζεται ή υπερβαίνει το περιεχόμενο και το δυναμικό εύρος των μικροσυστοιχιών (McPherson, 2009). Επίσης, η NGS προσέγγιση για την παρακολούθηση του μεταγραφώματος προσφέρει πολύ υψηλότερη ευαισθησία από ότι οι περισσότερες πλατφόρμες μικροσυστοιχιών καθώς τα θραύσματα των αλληλουχιών καλύπτουν ολόκληρες τις περιοχές των γονιδίων που είναι πολύ μεγαλύτερες από αυτές που καλύπτουν ειδικά σχεδιασμένοι ανιχνευτές. Η ανάλυση της γονιδιακής έκφρασης είναι δυνατή μέσα από τη χρήση του αντισώματος της RNA Πολυμεράσης II (RNAPII) σε συνδυασμό με ChIP για να τη σύλληψη των αλληλεπιδράσεων των μεταγραφικών συμπλόκων με περιοχές του DNA, ακολουθούμενο από βαθιά αλληλούχιση (Mortazavi et al., 2008). Μια ακόμα εφαρμογή αποτελεί η ανίχνευση ποικιλομορφιών αριθμού αντιγράφων (Copy Number Variation-CNVs). Αν και υπολογιστικές μέθοδοι βασισμένες σε δεδομένα μικροσυστοιχιών έχουν επιτυχώς χρησιμοποιηθεί για την αναγνώριση CNVs, η δύναμή τους είναι περιορισμένη καθώς το μέγεθος και η ανάλυση οποιονδήποτε προβλέψεων συσχετίζεται με την πυκνότητα των ανιχνευτών στη μικροσυστοιχία, η οποία με τη σειρά της περιορίζεται είτε από τη διακριτική ανάλυση της ίδιας της μικροσυστοιχίας, είτε από την πυκνότητα γνωστών περιοχών SNP (Medvedev et al., 2009). Έτσι, πρόσφατα, μεθοδολογίες βασισμένες σε αλληλούχιση ζευγαρωτών αλληλουχιών και σε συνδυασμό με τις πλατφόρμες NGS χρησιμοποιήθηκαν για την ανακάλυψη δομικών παραλλαγών (Bentley et al., 2008). Ωστόσο περιοχές ή και εξώνια πλούσια σε επαναλήψιμες ακολουθίες και με υψηλό ποσοστό βάσεων GC, προκειμένου να αλληλουχηθούν με

πληρότητα χρειάζονται επι του παρόντος την κλασσική τεχνική βασισμένη στη μέθοδο κατά Sanger (Valencia et al., 2012).

## 4. Συζήτηση – Συμπεράσματα

Τα τελευταία χρόνια, έχουν αναπτυχθεί και εξελιχθεί μέθοδοι αλληλούχισης που έχουν ξεπεράσει τεχνολογικά σε επίπεδο απόδοσης τις κλασσικές τεχνικές. Όλες αυτές οι μέθοδοι δεν βασίζονται στις κλασσικές πλατφόρμες αλληλούχισης κατά Sanger και περιλαμβάνουν τα συστήματα με τα εμπορικά ονόματα 454/Roche, Illumina, Applied Biosystems/Life Technologies, Dover, Helicos Biosciences, Ion Torrent και Ion Proton (Life technologies). Τα νέα αυτά συστήματα αλληλούχισης περιλαμβάνουν ένα κοινό πρότυπο αλληλούχισης που ονομάζεται κυκλική σειρά αλληλούχισης (Metzker, 2010). Αυτές οι πλατφόρμες επιτυγχάνουν σε χαμηλό κόστος μαζικές αλληλουχίσεις πολλών δειγμάτων ταυτόχρονα με ένα πολύ μεγάλο αριθμό διαβασμάτων ανά δείγμα. Βέβαια, προς το παρόν αυτή η πολύ υψηλή απόδοση επιτυγχάνεται με μειονεκτήματα όσο αφορά το μήκος και την ακρίβεια των διαβασμάτων, σε σύγκριση με τις κλασσικές μεθόδους, αλλά αυτά τα μειονεκτήματα υπερκαλύπτονται από τις εξαιρετικά πολυπληθείς αλληλοεπικαλύψεις των διαβασμάτων. Μάλιστα όταν υπάρχει γνωστό γονιδίωμα αναφοράς, όπως και στο ανθρώπινο αυτές οι νέες πλατφόρμες είναι άμεσα εφαρμόσιμες με πολύ μεγάλη απόδοση.

Τα τελευταία 5 χρόνια, οι υψηλής απόδοσης τεχνικές αλληλούχισης (high-throughput sequencing technologies) έχουν επιτύχει τον προσδιορισμό μεταλλάξεων σε νέα γονίδια που αφορούν σε ένα μεγάλο αριθμό γενετικών διαταραχών, όπως είναι για παράδειγμα τα σύνδρομα Sensenbrenner, Kabuki, και Miller (Gilissen et al. 2010; Ng et al., 2010). Η αλληλούχιση επόμενης γενιάς (**N**ext **G**eneration **S**equencing - NGS), έχει διευκολύνει σημαντικά σε ταχύτητα, ακρίβεια και με χαμηλό κόστος πολλές διαγνωστικές εφαρμογές. Επιπρόσθετα, οι τεχνικές των μικροσυστοιχιών (microarrays) και η αλυσιδωτή αντίδραση της πολυμεράσης (**P**olymerase **C**hain **R**eaction – PCR) (Albert et al., 2007; Gnirke et al., 2009; Hodges et al., 2007) μαζί με την αλληλούχιση επόμενης γενιάς NGS αποτελούν συμπληρωματικές τεχνολογίες πολύ σημαντικές για

πολλές σύγχρονες κλινικές εφαρμογές (Metzker, 2010; Voelkerding et al., 2010). Μάλιστα, τα τελευταία δύο χρόνια έχει αναπτυχθεί ένας μεγάλος αριθμός κλινικών εφαρμογών οι οποίες χρησιμοποιούν τη μέθοδο NGS προκειμένου να ανιχνευθούν μεταλλάξεις γονιδίων σε γενετικές ανωμαλίες, όπως είναι για παράδειγμα οι μυϊκές δυστροφίες και διάφορα είδη κώφωσης (Schrauwen et al., 2013; Sivakumaran et al., 2013; Valencia et al., 2013). Ένα άλλο χαρακτηριστικό παράδειγμα κλινικής εφαρμογής με την τεχνική της αλληλούχισης επόμενης γενιάς είναι η διάγνωση ανευπλοειδίας στα έμβρυα επιτόκων, η οποία εξαρτάται από το πατρικό DNA (Chiu et al. 2011; Ehrich et al. 2011; Palomaki et al., 2011). Όλα τα προαναφερθέντα αποτελούν μόνο λίγα παραδείγματα μοριακών κλινικών εφαρμογών της τεχνικής Next Generation Sequencing.

Τα επόμενα χρόνια, σε παγκόσμιο επίπεδο οι διαγνωστικές μέθοδοι μέσω μοριακών τεχνικών θα είναι πολύ σημαντικές για την πρόληψη και τις θεραπείες έναντι πολλών ασθενειών. Τα μοριακά γενετικά τεστ θα διευκολύνουν την ανίχνευση και τον χαρακτηρισμό ασθενειών και διαταραχών, καθώς θα ελέγχεται και η ανταπόκριση στις διάφορες αγωγές και θα προσδιορίζεται γενετικά η ευπάθεια στις νόσους για κάθε άτομο ξεχωριστά. Οι τεχνικές μαζικής αλληλούχισης είναι πιθανό να επικρατήσουν έναντι των τεχνικών αλληλούχισης υψηλής απόδοσης. Παρόλα αυτά, και άλλες προσεγγίσεις διερευνούνται προκειμένου να εξελιχθούν σε κλινικές μέθοδοι (Chan, 2005). Επίσης, ανεξάρτητα των μεθόδων, περιοχές ή και εξώνια πλούσια σε επαναλήψιμες ακολουθίες και με υψηλό ποσοστό βάσεων GC, προκειμένου να αλληλουχηθούν με πληρότητα θα χρειάζεται η κλασσική τεχνική βασισμένη στη μέθοδο κατά Sanger (Valencia et al., 2012).

Επίσης πρέπει να αναφερθεί και η σημασία της βιοπληροφορικής ως ένα ξεχωριστό τμήμα της αλληλούχισης επόμενης γενιάς, διότι είναι ένα ιδιαίτερο εργαλείο και μόνο μέσω της εξέλιξης των υπολογιστικών και αναλυτικών προγραμμάτων θα είναι εφικτή η συνδυασμένη εφαρμογή της τεχνικής στις κλινικές μεθόδους (Valencia et al., 2012).

Η αλληλούχιση επόμενης γενιάς έχει πολλά πλεονεκτήματα έναντι των κλασσικών μεθόδων αλληλούχισης κατά Sanger, με το πιο σημαντικό να αποτελεί το σχετικά χαμηλό κόστος σε μικρό χρονικό διάστημα (Wall et al., 2009). Ακόμη, μέσω αυτής της τεχνικής έχει βρεθεί ότι ένας μεγάλος αριθμός παρόμοιων φαινοτυπικά ασθενειών μπορεί να έχει πολλές διαφορετικές γενετικές αιτίες (Hoischen et al., 2010).



Αυτή η γενετική ετερογένεια η οποία παρουσιάζεται στις μυϊκές δυστροφίες, στις συγγενείς διαταραχές γλυκοζυλίωσης, στα διάφορα είδη απώλειας ακοής, μπορεί εύκολα να αντιμετωπιστεί εφαρμόζοντας την τεχνική NGS, δηλαδή με αλληλούχιση όλων των σχετικών γονιδίων με τους ειδικούς φαινοτύπους που παρουσιάζονται (Valencia et al., 2012). Μάλιστα, στην περίπτωση των συγγενών διαταραχών γλυκοζυλίωσης, η τεχνική NGS αποτελεί ένα χρονικά και οικονομικά αποδοτικό εργαλείο με το οποίο ανιχνεύονται μεταλλάξεις, που σχετίζονται με μεταβολικές ασθένειες σε διαφορετικά γονίδια του ίδιου βιοχημικού μονοπατιού. Επίσης, αξιοσημείωτα είναι τα πλεονεκτήματα στο πεδίο της γενετικής του καρκίνου, όπου πολύ γρήγορα μπορούν να εξαχθούν συμπεράσματα για τη διάγνωση και το είδος του καρκινικού όγκου (Stratton et al. 2009). Ακόμη, μία άλλη εφαρμογή αποτελεί η προγενετική διάγνωση ανευπλοειδίων και τρισωμιών στα χρωμοσώματα 13, 18 και 21 (Chiu et al., 2008).

Επιπρόσθετα, αξιοσημείωτο είναι το θέμα της βιοηθικής προσέγγισης της αλληλούχισης επόμενης γενιάς (Pinxten and Howard, 2014). Τίθενται πολύ σημαντικά ερωτήματα που η απάντησή τους μπορεί να βρεθεί μόνο από διεπιστημονική προσέγγιση. Για παράδειγμα, υπάρχουν περιπτώσεις στις οποίες το ιατρικό ζήτημα πιθανόν να μπορεί να απαντηθεί από μια μικρής κλίμακας εξέταση του γονιδιώματος, οπότε σε αυτές τις περιπτώσεις δεν θα είναι υποχρεωτική η χρησιμοποίηση της τεχνικής NGS, και αυτό θα κρίνεται από εξειδικευμένους επιστήμονες. Ένα άλλο ερώτημα είναι αν οι ασθενείς θα ενημερώνονται όταν τα αποτελέσματα δίνουν μόνο ενδείξεις με αμφίβολη κλινική σημασία ή πολύ σοβαρές ασθένειες που δεν μπορούν να προληφθούν ή να αντιμετωπιστούν. Σημαντική επίσης είναι η αποθήκευση των αποτελεσμάτων και η ενδεχόμενη χρησιμοποίησή τους από τρίτους εκτός ιατρικών υποθέσεων, όπως μπορεί να είναι οι ασφαλιστικές υπηρεσίες. Πλέον γίνεται κατανοητό πως το γενετικό προφίλ κάθε ανθρώπου μπορεί να δώσει σημαντικές πληροφορίες για την υγεία του και την κατάσταση του οργανισμού, οπότε οι πληροφορίες του γονιδιώματος πρέπει να αποθηκεύονται και να χειρίζονται με τέτοιο τρόπο έτσι, ώστε να τηρείται απόλυτα το ιατρικό απόρρητο και αναλόγως των ιατρικών ερωτημάτων που τίθενται να δίνονται οι κατάλληλες απαντήσεις, είτε στον ίδιο τον ασθενή, είτε για την αντιμετώπιση των ενδεχόμενων παθολογικών καταστάσεων.

## 5. Βιβλιογραφία

Audo, I., Bujakowska, K., Orhan, E., Poloschek, C. M., Defoort-Dhellemmes, S., Drumare, I., & Sahel, J. A. (2012). Whole-Exome Sequencing Identifies Mutations in *GPR179* Leading to Autosomal-Recessive Complete Congenital Stationary Night Blindness. *The American Journal of Human Genetics*, *90*(2), 321-330.

Acland, A., Agarwala, R., Barrett, T., Beck, J., Benson, D. A., Bollin, C. & Slotta, D. (2013). Database resources of the national center for biotechnology information. *Nucleic acids research*, *41*(D1), D8-D20.

Adessi, C., Matton, G., Ayala, G., Turcatti, G., Mermod, J. J., Mayer, P., & Kawashima, E. (2000). Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic acids research*, *28*(20), e87-e87.

Albert, T. J., Molla, M. N., Muzny, D. M., Nazareth, L., Wheeler, D., Song, X., & Gibbs, R. A. (2007). Direct selection of human genomic loci by microarray hybridization. *Nature methods*, *4*(11), 903-905.

Bell, J. T., & Spector, T. D. (2011). A twin approach to unraveling epigenetics. *Trends in Genetics*, *27*(3), 116-125.

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G. & Anastasi, C. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, *456*(7218), 53-59.

Berger, M. F., Lawrence, M. S., Demichelis, F., Drier, Y., Cibulskis, K., Sivachenko, A. Y. & Garraway, L. A. (2011). The genomic complexity of primary human prostate cancer. *Nature*, *470*(7333), 214-220.

Beroukhim, R., Mermel, C. H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J. & Letai, A. (2010). The landscape of somatic copy-number alteration across human cancers. *Nature*, *463*(7283), 899-905.

Blow, N. (2008). Cell imaging: new ways to see a smaller world. *Nature*, *456*(7223), 825-828.

Calvo, S. E., Compton, A. G., Hershman, S. G., Lim, S. C., Lieber, D. S., Tucker, E. J. & Mootha, V. K. (2012). Molecular diagnosis of infantile mitochondrial disease with targeted next-generation sequencing. *Science translational medicine*, *4*(118), 118ra10-118ra10.

Casey, J. P., Magalhaes, T., Conroy, J. M., Regan, R., Shah, N., Anney, R., ... & Le Couteur, A. (2012). A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Human genetics*, *131*(4), 565-579.

Chen, A., Xu, J., & Johnson, A. C. (2006). Curcumin inhibits human colon cancer cell growth by suppressing gene expression of epidermal growth factor receptor through reducing the activity of the transcription factor Egr-1. *Oncogene*, *25*(2), 278-287.

Chiu, R. W., Chan, K. A., Gao, Y., Lau, V. Y., Zheng, W., Leung, T. Y. & Lo, Y. D. (2008). Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proceedings of the National Academy of Sciences*, *105*(51), 20458-20463.

Chiu, R. W., Akolekar, R., Zheng, Y. W., Leung, T. Y., Sun, H., Chan, K. A. & Lo, Y. D. (2011). Non-invasive prenatal assessment of trisomy 21 by multiplexed maternal plasma DNA sequencing: large scale validity study. *BMJ: British Medical Journal*, *342*.

Daiger, S. P., Sullivan, L. S., Bowne, S. J., Birch, D. G., Heckenlively, J. R., Pierce, E. A., & Weinstock, G. M. (2010). Targeted high-throughput DNA sequencing for gene discovery in retinitis pigmentosa. In *Retinal Degenerative Diseases* (pp. 325-331). Springer New York.

Das, P. M., & Singal, R. (2004). DNA methylation and cancer. *Journal of Clinical Oncology*, 22(22), 4632-4642.

Desai, A. N., & Jere, A. (2012). Next-generation sequencing: ready for the clinics?. *Clinical genetics*, 81(6), 503-510.

Dressman, D., Yan, H., Traverso, G., Kinzler, K. W., & Vogelstein, B. (2003). Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proceedings of the National Academy of Sciences*, 100(15), 8817-8822.

Dunham, I., Hunt, A. R., Collins, J. E., Bruskiewich, R., Beare, D. M., Clamp, M., ... & Grafham, D. (1999). The DNA sequence of human chromosome 22. *Nature*, 402(6761), 489-495.

Ehrich, M., Deciu, C., Zwiefelhofer, T., Tynan, J. A., Cagasan, L., Tim, R., ... & van den Boom, D. (2011). Noninvasive detection of fetal trisomy 21 by sequencing of DNA in maternal blood: a study in a clinical setting. *American journal of obstetrics and gynecology*, 204(3), 205-e1.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G. & Zhao, P. (2009). Real-time DNA sequencing from single polymerase molecules. *Science*, 323(5910), 133-138.

Fedurco, M., Romieu, A., Williams, S., Lawrence, I., & Turcatti, G. (2006). BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic acids research*, 34(3), e22-e22.

Forbes, S. A., Bindal, N., Bamford, S., Cole, C., Kok, C. Y., Beare, D., & Futreal, P. A. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic acids research*, 39(suppl 1), D945-D950.

Frueh, F. W., Amur, S., Mummaneni, P., Epstein, R. S., Aubert, R. E., DeLuca, T. M. & Lesko, L. J. (2008). Pharmacogenomic biomarker information in drug labels approved by the United States food and drug administration: prevalence of related drug use. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, 28(8), 992-998.

Futreal, P. A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., & Miki, Y. (1994). BRCA1 mutations in primary breast and ovarian carcinomas. *Science*, 266(5182), 120-122.

Garraway, L. A. (2013). Genomics-driven oncology: framework for an emerging paradigm. *Journal of Clinical Oncology*, 31(15), 1806-1814.

Ghosh, S., Krux, F., Binder, V., Gombert, M., Niehues, T., Feyen, O. and Borkhardt, A. (2012). Array-Based Sequence Capture and Next-Generation Sequencing for the Identification of Primary Immunodeficiencies. *Scandinavian journal of immunology*, 75(3), 350-354.

Garraway, L. A., Verweij, J., & Ballman, K. V. (2013). Precision oncology: an overview. *Journal of Clinical Oncology*, 31(15), 1803-1805.

Gilissen, C., Arts, H. H., Hoischen, A., Spruijt, L., Mans, D. A., Arts, P. & Brunner, H. G. (2010). Exome Sequencing Identifies WDR35 Variants Involved in Sensenbrenner Syndrome. *The American Journal of Human Genetics*, 87(3), 418-423.

Gnirke, A., Melnikov, A., Maguire, J., Rogov, P., LeProust, E. M., Brockman, W. & Nusbaum, C. (2009). Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nature biotechnology*, 27(2), 182-189.

Gong, Y., McDonough, C. W., Wang, Z., Hou, W., Cooper-DeHoff, R. M., Langae, T. Y. & Johnson, J. A. (2012). Hypertension Susceptibility Loci and Blood Pressure Response to Antihypertensives Results From the Pharmacogenomic Evaluation of

Antihypertensive Responses Study. *Circulation: Cardiovascular Genetics*, 5(6), 686-691.

Guo, R. J., Funakoshi, S., Lee, H. H., Kong, J., & Lynch, J. P. (2010). The intestine-specific transcription factor Cdx2 inhibits  $\beta$ -catenin/TCF transcriptional activity by disrupting the  $\beta$ -catenin–TCF protein complex. *Carcinogenesis*, 31(2), 159-166.

Harris, T. D., Buzby, P. R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., ... & Xie, Z. (2008). Single-molecule DNA sequencing of a viral genome. *Science*, 320(5872), 106-109.

Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M. N., Smith, S. W. & McCombie, W. R. (2007). Genome-wide in situ exon capture for selective resequencing. *Nature genetics*, 39(12), 1522-1527.

Hoischen, A., Gilissen, C., Arts, P., Wieskamp, N., van der Vliet, W., Vermeer, S., & Veltman, J. A. (2010). Massively parallel sequencing of ataxia genes after array-based enrichment. *Human mutation*, 31(4), 494-499.

Housby, J. N., & Southern, E. M. (1998). Fidelity of DNA ligation: a novel experimental approach based on the polymerisation of libraries of oligonucleotides. *Nucleic acids research*, 26(18), 4259-4266.

Hutchison, C. A. (2007). DNA sequencing: bench to bedside and beyond. *Nucleic acids research*, 35(18), 6227-6237.

Janipalli, C. S., Kumar, M. V. K., Vinay, D. G., Sandeep, M. N., Bhaskar, S., Kulkarni, S. R. & Chandak, G. R. (2012). Analysis of 32 common susceptibility genetic variants and their combined effect in predicting risk of Type 2 diabetes and related traits in Indians. *Diabetic medicine*, 29(1), 121-127.

Jones, M. A., Bhide, S., Chin, E., Ng, B. G., Rhodenizer, D., Zhang, V. W., & Hegde, M. R. (2011). Targeted polymerase chain reaction-based enrichment and next generation

sequencing for diagnostic testing of congenital disorders of glycosylation. *Genetics in Medicine*, 13(11), 921-932.

Jones, M. L., Murden, S. L., Bem, D., Mundell, S. J., Gissen, P., Daly, M. E. & Mumford, A. D. (2012). Rapid genetic diagnosis of heritable platelet function disorders with next-generation sequencing: proof-of-principle with Hermansky–Pudlak syndrome. *Journal of Thrombosis and Haemostasis*, 10(2), 306-309.

Lander, E. S., & Waterman, M. S. (1988). Genomic mapping by fingerprinting random clones: a mathematical analysis. *Genomics*, 2(3), 231-239.

Levy, S., Sutton, G., Ng, P. C., Feuk, L., Halpern, A. L., Walenz, B. P., & Venter, J. C. (2007). The diploid genome sequence of an individual human. *PLoS biology*, 5(10), e254.

Link, D. C., Schuettpelez, L. G., Shen, D., Wang, J., Walter, M. J., Kulkarni, S., & Mardis, E. R. (2011). Identification of a novel TP53 cancer susceptibility mutation through whole-genome sequencing of a patient with therapy-related AML. *Jama*, 305(15), 1568-1576.

McCormick, D. A., & Contreras, D. (2001). On the cellular and network bases of epileptic seizures. *Annual Review of Physiology*, 63(1), 815-846.

McNamara, J. O. (1994). Cellular and molecular basis of epilepsy. *The Journal of Neuroscience*, 14(6), 3413-3425.

MacConaill, L. E. (2013). Existing and emerging technologies for tumor genomic profiling. *Journal of Clinical Oncology*, 31(15), 1815-1824.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., & Volkmer, G. A. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), 376-380.

Marrone M, Filipski KK, Gillanders EM, Schully SD, Freedman AN. Multi-marker Solid Tumor Panels Using Next-generation Sequencing to Direct Molecularly Targeted Therapies. *PLOS Currents Evidence on Genomic Tests*. 2014 May 27

Maxam, A. M., & Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods in enzymology*, 65(1), 499.

McClellan, J., & King, M. C. (2010). Genetic heterogeneity in human disease. *Cell*, 141(2), 210-217.

McPherson, J. D. (2009). Next-generation gap. *Nature Methods*, 6, S2-S5.

Meder, B., Haas, J., Keller, A., Heid, C., Just, S., Borries, A., ... & Rottbauer, W. (2011). Targeted next-generation sequencing for the molecular genetic diagnostics of cardiomyopathies. *Circulation: Cardiovascular Genetics*, 4(2), 110-122.

Medvedev, P., Stanciu, M., & Brudno, M. (2009). Computational methods for discovering structural variation with next-generation sequencing. *Nature methods*, 6, S13-S20.

Metzker, M. L. (2009). Sequencing in real time. *Nature biotechnology*, 27(2), 150.

Metzker, M. L. (2010). Sequencing technologies—the next generation. *Nature Reviews Genetics*, 11(1), 31-46.

Mierendorf, R. C., & Pfeffer, D. (1987). Direct sequencing of denatured plasmid DNA. *Methods in enzymology*, 152, 556.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), 621-628.



Murray, J. C., Buetow, K. H., Weber, J. L., Ludwigsen, S., Scherpbier-Heddema, T., Manion, F., & Duyk, G. M. (1994). A comprehensive human linkage map with centimorgan density. Cooperative Human Linkage Center (CHLC). *Science*, 265(5181), 2049-2054.

Ng, S. B., Bigham, A. W., Buckingham, K. J., Hannibal, M. C., McMillin, M. J., Gildersleeve, H. I. & Shendure, J. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nature genetics*, 42(9), 790-793.

Onoufriadis, A., Shoemark, A., Schmidts, M., Patel, M., Jimenez, G., Liu, H., ... & Mitchison, H. M. (2014). Targeted NGS gene panel identifies mutations in RSPH1 causing primary ciliary dyskinesia and a common mechanism for ciliary central-pair agenesis due to radial spoke defects. *Human molecular genetics*, ddu046.

Onoufriadis, A., Shoemark, A., Munye, M. M., James, C. T., Schmidts, M., Patel, M., ... & Mitchison, H. M. (2014). Combined exome and whole-genome sequencing identifies mutations in ARMC4 as a cause of primary ciliary dyskinesia with defects in the outer dynein arm. *Journal of medical genetics*, 51(1), 61-67.

Palomaki, G. E., Kloza, E. M., Lambert-Messerlian, G. M., Haddow, J. E., Neveux, L. M., Ehrich, M. & Canick, J. A. (2011). DNA sequencing of maternal plasma to detect Down syndrome: an international clinical validation study. *Genetics in medicine*, 13(11), 913-920.

Petrova, T. V., Bono, P., Holnthoner, W., Chesnes, J., Pytowski, B., Sihto, H., ... & Alitalo, K. (2008). VEGFR-3 expression is restricted to blood and lymphatic vessels in solid tumors. *Cancer cell*, 13(6), 554-556.

Pinxten, W., & Howard, H. C. (2014). Ethical issues raised by whole genome sequencing. *Best Practice & Research Clinical Gastroenterology*, 28(2), 269-279.

Privitera, M. (2011). Current challenges in the management of epilepsy. *Am. J. Manag. Care*, 17, S195-203.

- Rothberg, J. M., & Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nature biotechnology*, 26(10), 1117-1124.
- Sanger, F. (1988). Sequences, sequences, and sequences. *Annual review of biochemistry*, 57(1), 1-29.
- Sarda, S., & Hannehalli, S. (2014). Next-Generation Sequencing and Epigenomics Research: A Hammer in Search of Nails. *Genomics & informatics*, 12(1), 2-11.
- Shendure, J. A., Porreca, G. J., Church, G. M., Gardner, A. F., Hendrickson, C. L., Kieleczawa, J., & Slatko, B. E. (2011). Overview of DNA sequencing strategies. *Current Protocols in Molecular Biology*, 7-1.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463-5467.
- Sanger, F., & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of molecular biology*, 94(3), 441-448.
- Schrauwen, I., Sommen, M., Corneveaux, J. J., Reiman, R. A., Hackett, N. J., Claes, C. & Huentelman, M. J. (2013). A sensitive and specific diagnostic test for hearing loss using a microdroplet PCR-based approach and next generation sequencing. *American Journal of Medical Genetics Part A*, 161(1), 145-152.
- Sinsheimer, R. L. (2006). To reveal the genomes. *American journal of human genetics*, 79(2), 194.
- Sivakumaran, T. A., Husami, A., Kissell, D., Zhang, W., Keddache, M., Black, A. P., & Zhang, K. (2013). Performance evaluation of the next-generation sequencing approach

for molecular diagnosis of hereditary hearing loss. *Otolaryngology--Head and Neck Surgery*, 148(6), 1007-1016.

Stratton, M. R., Campbell, P. J., & Futreal, P. A. (2009). The cancer genome. *Nature*, 458(7239), 719-724.

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics*, 9(6), 465-476.

Tabor, S., & Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proceedings of the National Academy of Sciences*, 84(14), 4767-4771.

Tettelin, H., & Feldblyum, T. (2009). Bacterial genome sequencing. In *Molecular Epidemiology of Microorganisms* (pp. 231-247). Humana Press.

Valencia, C. A., Ankala, A., Rhodenizer, D., Bhide, S., Littlejohn, M. R., Keong, L. M., & Hegde, M. (2013). Comprehensive mutation analysis for congenital muscular dystrophy: a clinical PCR-based enrichment and next-generation sequencing panel. *PloS one*, 8(1), e53083.

Valouev, A., Johnson, D. S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., & Sidow, A. (2008). Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nature methods*, 5(9), 829-834.

Van Allen, E. M., Wagle, N., & Levy, M. A. (2013). Clinical analysis and interpretation of cancer genome data. *Journal of Clinical Oncology*, 31(15), 1825-1833.

Vasta, V., Ng, S. B., Turner, E. H., Shendure, J., & Hahn, S. H. (2009). Next generation sequence analysis for mitochondrial disorders. *Genome Med*, 1(10), 100.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., & Beasley, E. (2001). The sequence of the human genome. *science*, 291(5507), 1304-1351.

Voelkerding, K. V., Dames, S., & Durtschi, J. D. (2010). Next generation sequencing for clinical diagnostics-principles and application to targeted resequencing for hypertrophic cardiomyopathy: a paper from the 2009 William Beaumont Hospital Symposium on Molecular Pathology. *The Journal of Molecular Diagnostics*, 12(5), 539-551.

Waddington, C. H. (1953, January). Epigenetics and evolution. In *Symp. Soc. Exp. Biol* (Vol. 7, pp. 186-199).

Wall, P. K., Leebens-Mack, J., Chanderbali, A. S., Barakat, A., Wolcott, E., Liang, H. & Altman, N. (2009). Comparison of next generation sequencing technologies for transcriptome characterization. *Bmc Genomics*, 10(1), 347.

Wang, J., Gotway, G., Pascual, J. M., & Park, J. Y. (2014). Diagnostic Yield of Clinical Next-Generation Sequencing Panels for Epilepsy. *JAMA neurology*, 71(5), 650-651.

Weber, M., & Schübeler, D. (2007). Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Current opinion in cell biology*, 19(3), 273-280.

Wong, L. J. C. (2013). *Next Generation Sequencing: Translation to Clinical Diagnostics*. Springer.

Wu, T. J., Shamsaddini, A., Pan, Y., Smith, K., Crichton, D. J., Simonyan, V., & Mazumder, R. (2014). A framework for organizing cancer-related variations from existing databases, publications and NGS data using a High-performance Integrated Virtual Environment (HIVE). *Database*, 2014, bau022.

Zogopoulos, G., Ha, K. C., Naqib, F., Moore, S., Kim, H., Montpetit, A., & Gallinger, S. (2007). Germ-line DNA copy number variation frequencies in a large North American population. *Human genetics*, 122(3-4), 345-353.