

ΒΑΣΙΚΕΣ ΑΡΧΕΣ ΕΠΟΜΕΝΗΣ ΓΕΝΙΑΣ ΑΛΛΗΛΟΥΧΙΣΗΣ (NEXT GENERATION SEQUENCING, NGS)

ΚΟΥΝΙΑΚΗ ΔΙΑΜΑΝΤΩ
ΒΙΟΛΟΓΟΣ
ΤΜΗΜΑ ΑΝΟΣΟΛΟΓΙΑΣ-ΙΣΤΟΣΥΜΒΑΤΟΤΗΤΑΣ
Γ.Ν.Α. «Ο ΕΥΑΓΓΕΛΙΣΜΟΣ»

24^ο Ετήσιο Σεμινάριο Συνεχιζόμενης Ιατρικής Εκπαίδευσης
Νοσοκομείου «ο Ευαγγελισμός»
Αθήνα, 18-22 Φεβρουαρίου 2019

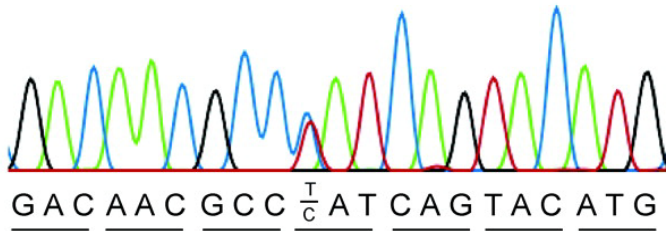


Δεν υπάρχει σύγκρουση συμφερόντων
με τις παρακάτω χορηγούς εταιρείες:

PFIZER, JANSSEN ONCOLOGY, SOFMEDICA,
NOVARTIS, ABBVIE, MSD, WINMEDICA,
GENESIS, ROCHE, TAKEDA, ASTELLAS,
AMGEN, ANGELINI, ANTISEL, SERVIER,
BRISTOL-MYERS SQUIBB, ABBOTT, GILEAD,
SANDOZ, BIANEΞ, RONTIS, MAVROGENIS,
AENORASIS, SPECIFAR, KARYO

- **1^η γενιά ανάλυσης νουκλεοτιδικής αλληλουχίας (1967) :**
 - κατά Sanger
 - αποτιμήθηκε ως η πιο ολοκληρωμένη μέθοδος αλληλούχησης
 - σαφώς καθορισμένη χημεία, υψηλής ακρίβειας
 - αποκωδικοποίηση του ανθρωπίνου γονιδιώματος (10ετία / $\sim 70 \times 10^6$ \$)
- **2^η γενιά ανάλυσης νουκλεοτιδικής αλληλουχίας (2005):**
 - χρόνος αλληλούχησης ενός μικρού γονιδιώματος 1 ημέρα
 - >5 ανθρώπινα γονιδιώματα μπορούν να αναλυθούν σε ένα run
 - κόστος < \$ 5000 / γονιδίωμα
- **3^η γενιά ανάλυσης νουκλεοτιδικής αλληλουχίας (2011):**
 - αλληλούχηση ολόκληρου-μεταγραφήματος
 - χωρίς ανάγκη συναρμολόγησης

Sanger Sequencing (~115kb/ημέρα, αδυναμία φάσης των ετεροζυγωτικών αλληλουχιών, περιορισμένα μήκη ανάγνωσης / 500 bp - 1 kb)



2x coverage

LOD 10-20%

Υψηλή ανάλυση → αύξηση κόστους

NGS (~80Mb/run, μήκη ανάγνωσης 150-700 bp)

```
AAAACCAGAGTCTAGCACCTTCTCATCAGGAGCAG
AAACCAGAGTCTAGCACCTTCTCATCAGGAGCAAC
AACCAGAGTCTAGCACCTTCTCATCAGGAGCAACG
ACCAGAGTCTAGCACCTTCTCATCAGGAGCAACGT
ACCAGAGTCTAGCACCTTCTCATCAGGAGCAGCGT
CCAGAGTCTAGCACCTTCTCATCAGGAGCAACGTC
GAGTCTAGCACCTTCTCATCAGGAGCAACGTCTGC
CTAGCACCTTCTCATCAGGAGCAGCGTCTGCCTTC
TAGCACCTTCTCATCAGGAGCAACGTCTGCCTTCG
AGCACCTTCTCATCAGGAGCAACGTCTGCCTTCGC
CCCTTCTCATCAGGAGCAGCGTCTGCCTTCGCTAG
ACCTTCTCATCAGGAGCAACGTCTGCCTTCGCTAG
CTTCTCATCAGGAGCAACGTCTGCCTTCGCTAGGC
ATCAGGAGCAACGTCTGCCTTCGCTAGGCTGACAT
ATCAGGAGCAACGTCTGCCTTCGCTAGGCTGACAT
TCAGGAGCAACGTCTGCCTTCGCTAGGCTGACAT
```

~30x

20 – 10.000x coverage

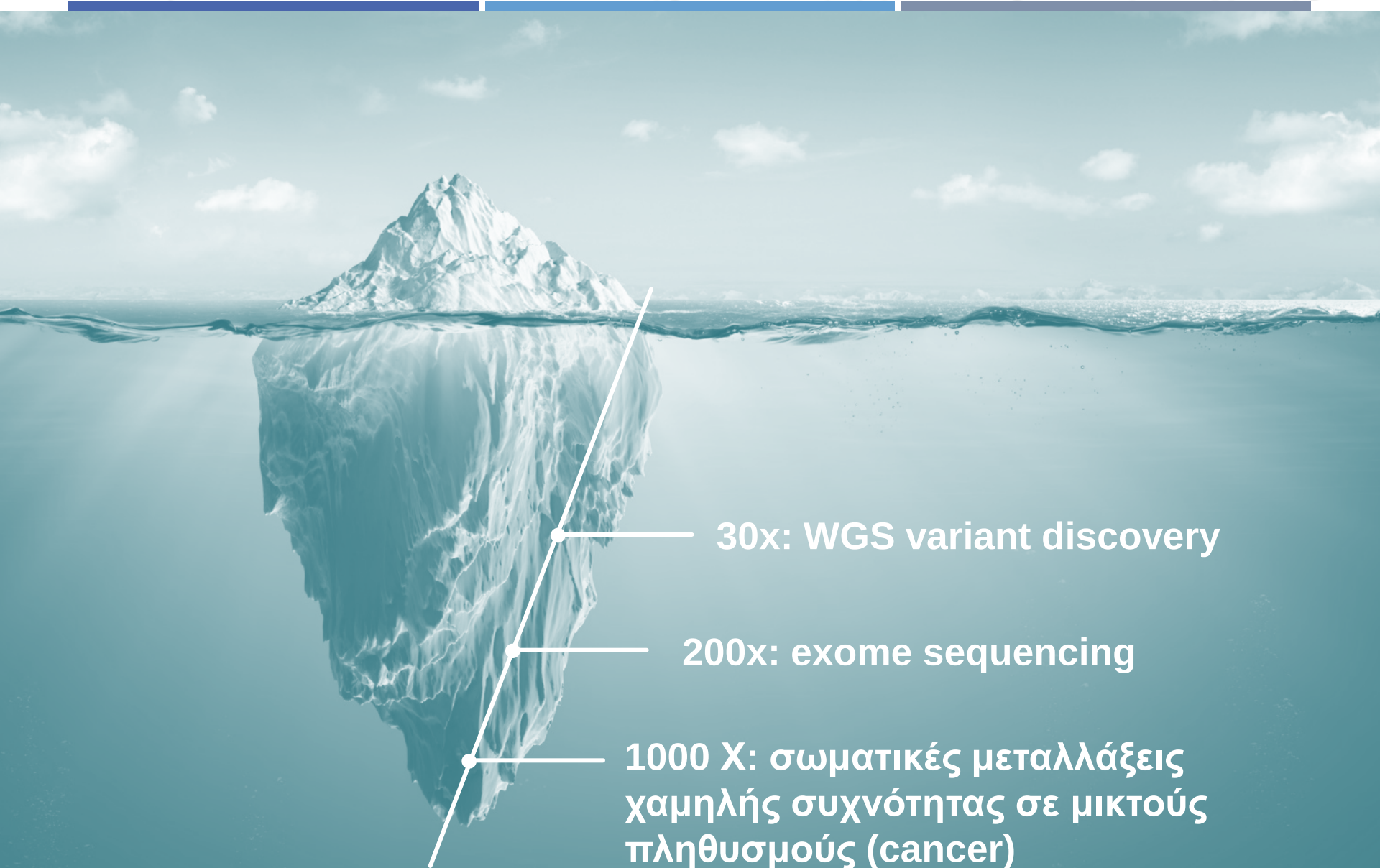
Massively parallel

1 run= “Millions Sanger αντιδράσεις”

LOD < 5%

A
G

Ευελιξία στο επίπεδο ανάλυσης-Επιλογή βάθους αλληλούχισης



High-Throughput DNA Sequencing, HTS

προϋποθέτουν PCR ενίσχυση

2nd Generation / Deep / Next Generation Sequencing-NGS (**150-700bp**)

- o **Roche**/454 GS FLX Titanium/Junior (2005)
- o **SOLiD** (Applied Biosystems, ABI) (2007)
- o Miseq, HiSeq, NextSeq **Illumina**/Solexa (2011)
- o PGM, S5 **Ion Torrent** (Life Technologies) (2015)

εμπορικά διαθέσιμες /
κλινικές εφαρμογές

η αλληλούχιση γίνεται άμεσα

3rd Generation / Next-Next Generation Sequencing (**>10 Kbp**)

- o Heliscope / Helicos (2010)
- o Pacific Biosciences (SMRT) (2011)
- o MinION (Oxford Nanopore) (2014)

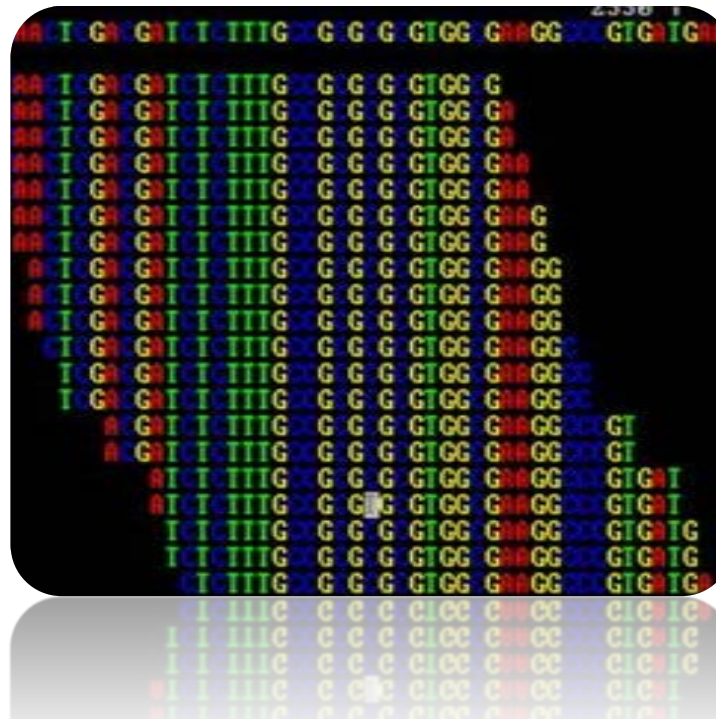


κλωνική ενίσχυση

απομόνωση ενός **μεμονωμένου μορίου DNA** και ενίσχυσή του, ώστε να παραχθεί ένας κλωνικός πληθυσμός από πανομοιότυπα γονιδιακά τμήματα

υψηλό επίπεδο παραλληλισμού «massively parallel»

ταυτόχρονη ανάλυση εκατοντάδων χιλιάδων πανομοιότυπων τμημάτων DNA



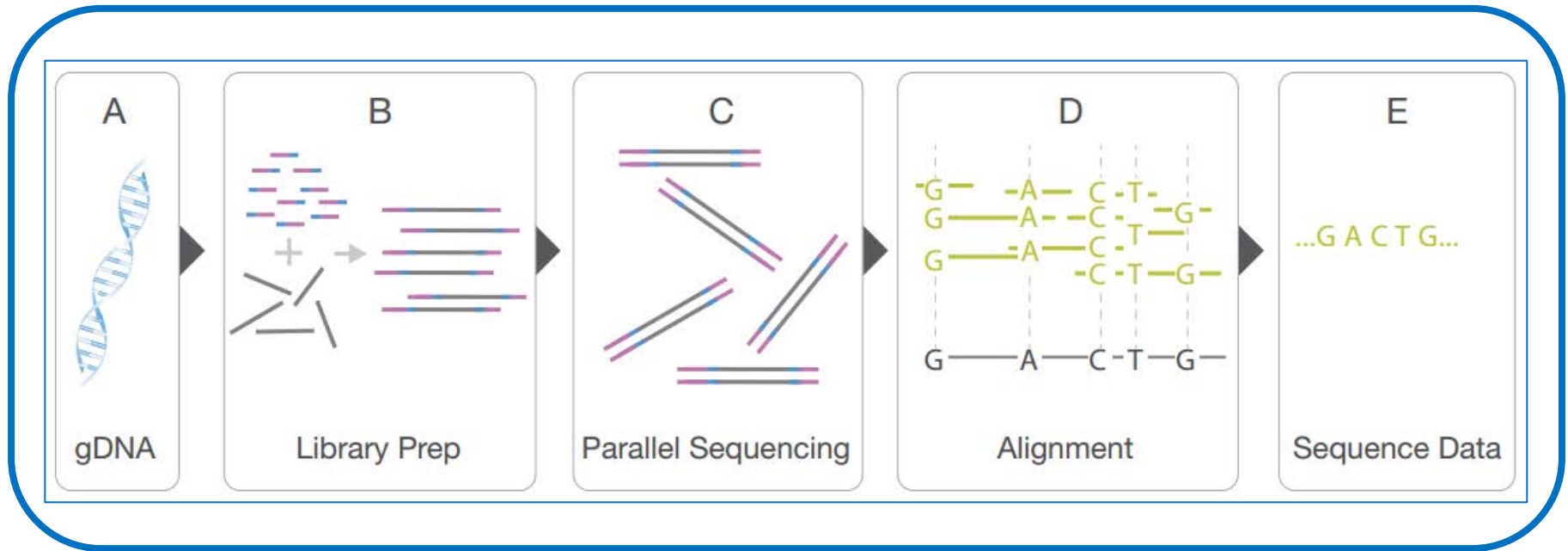
ο στοχευμένη αλληλούχηση εξονίου βασιζόμενη

σε amplicon (η πλέον χρησιμοποιούμενη διαδρομή)

ο **cDNA στοχευμένη αλληλούχηση μεταγράφων**

- επιτρέπουν την λήψη ευρύτερων πληροφοριών
- ταυτοποίηση νέων αλληλομόρφων
- επέκταση των αλληλουχιών αναφοράς σε γενετικές περιοχές που δεν έχουν ακόμη περιγραφεί

2^{ης} γενιάς μεθοδολογίες αλληλούχισης



α) εκχύλιση g-DNA/cDNA/RNA

β) παρασκευή βιβλιοθήκης

γ) κλωνική ενίσχυση

δ) συναρμολόγηση / ευθυγράμμιση

ε) ανάλυση των γενετικών δεδομένων

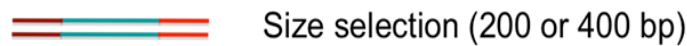


Δύο εργάσιμες ημέρες



Αυτοματοποιημένα

Genomic DNA / RNA



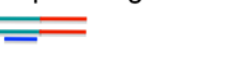
Biotinylated RNA oligos designed for selected genomic regions

Streptavidin magnetic beads

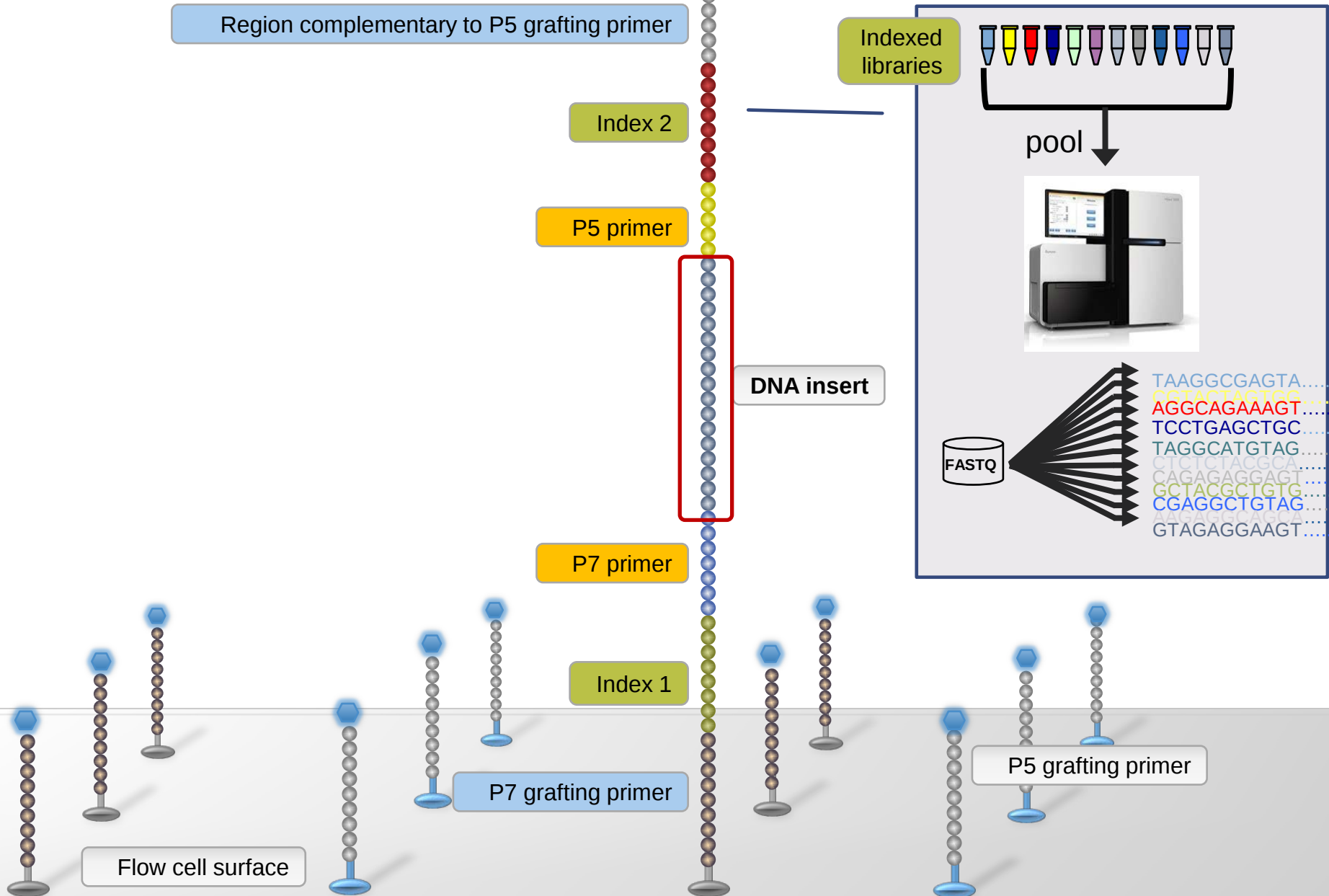
RNA Digestion, Wash

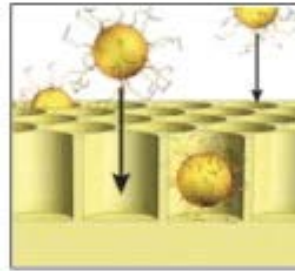
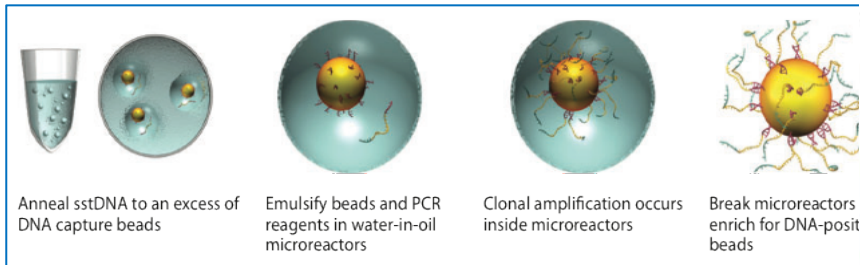
PCR amplification

Single end sequencing or paired-end sequencing



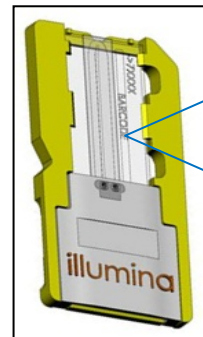
Προετοιμασία βιβλιοθήκης



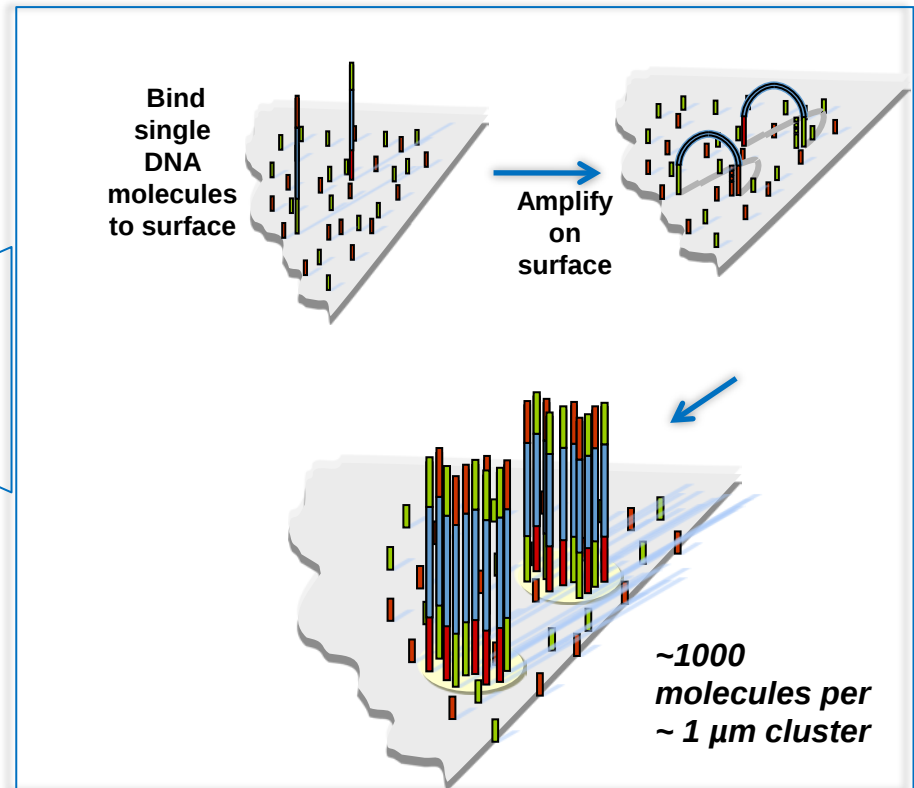


Emulsion PCR :
Roche
Ion Torrent

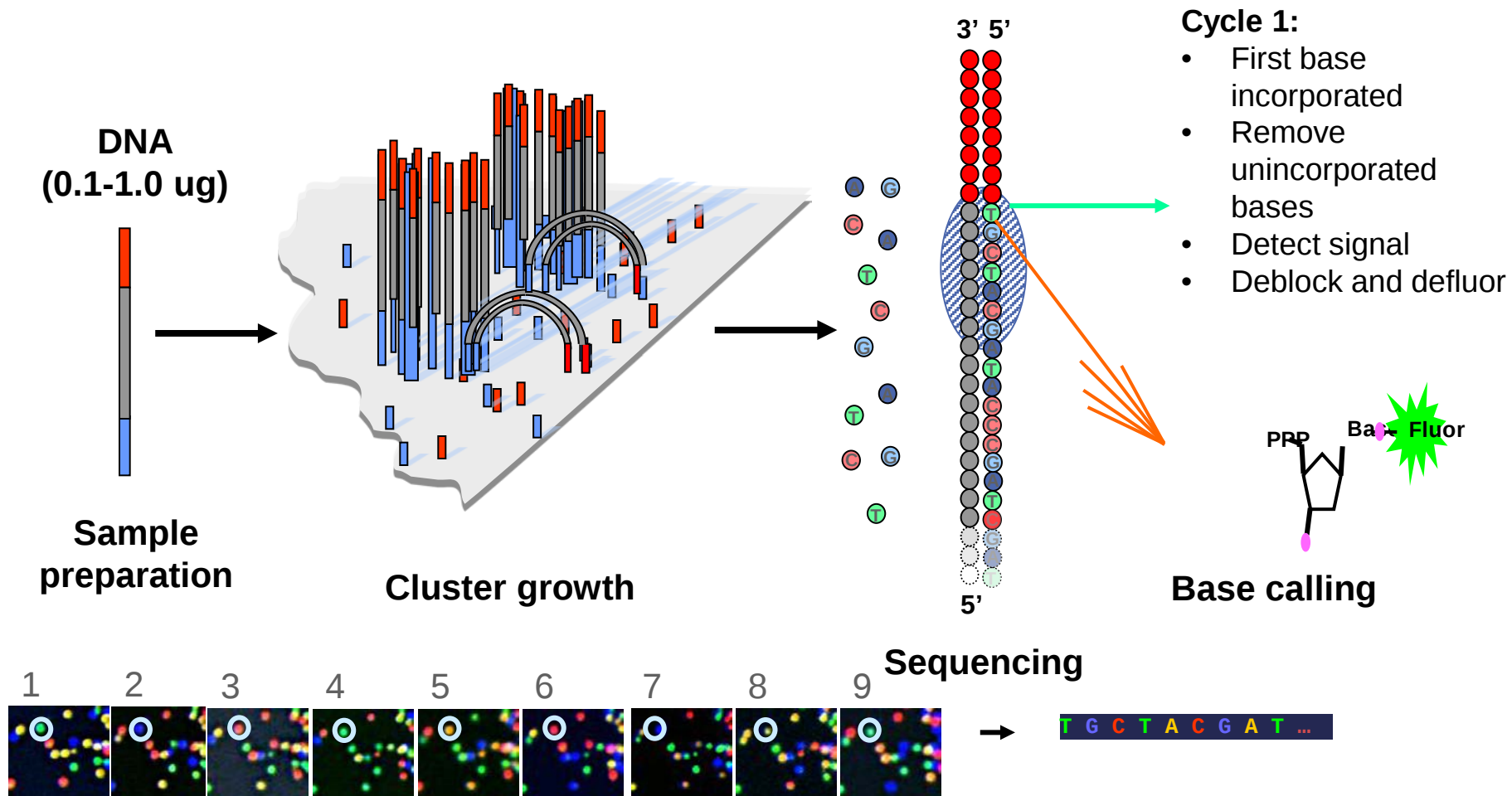
Bridge PCR :
Illumina
Life echnologies/SOLiD



κυπελίδα
ροής

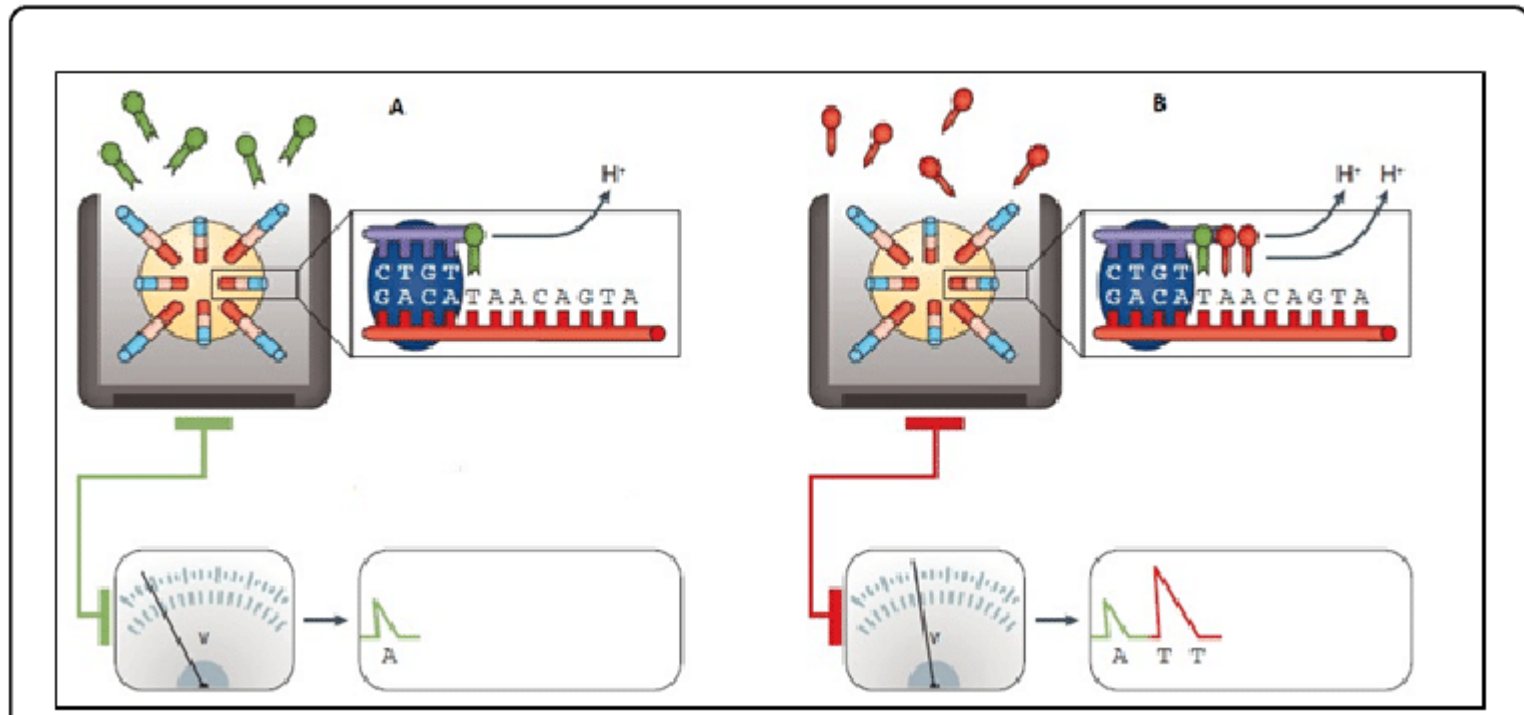


Sequencing by Synthesis with reversible terminator σε κυψελίδα ροής (Illumina, Life Technologies/SOLiD)



αλληλούχιση με σύνθεση, με βάση το φθορισμό των ανάστροφα τροποποιημένων με φθορίζουσα ουσία νουκλεοτιδίων τερματισμού

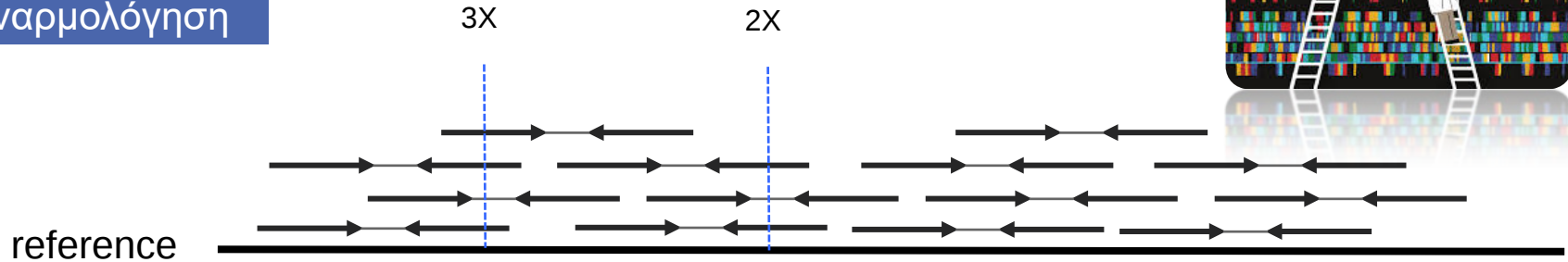
Sequencing by ligation σε σφαιρίδια (Ion Torrent)



- ο Sequencing by ligation χρησιμοποιώντας μη τροποποιημένα dNTPs
- ο η **απελευθέρωση ιόντων H^{++} μεταβάλλει το τοπικό pH**: ανιχνεύεται και μεταφράζεται σε ψηφιακό σήμα
- ο PCR αντίδραση γαλακτώματος / κλωνική ενίσχυση πάνω σε σφαιρίδια

Βιοπληροφορική ανάλυση 2^{ης} γενιάς αλληλούχησης

Ευθυγράμμιση
Συναρμολόγηση



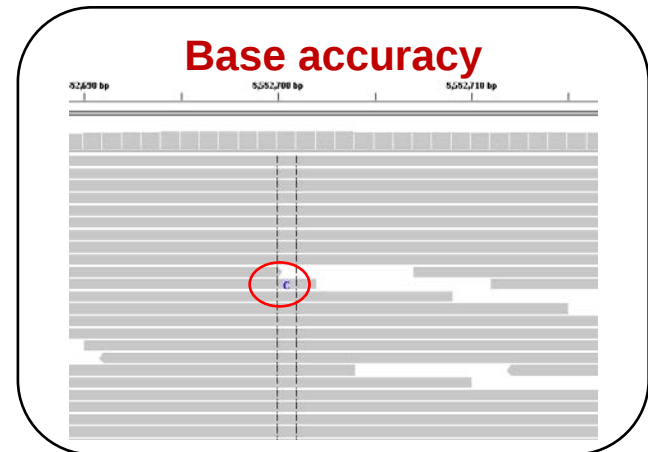
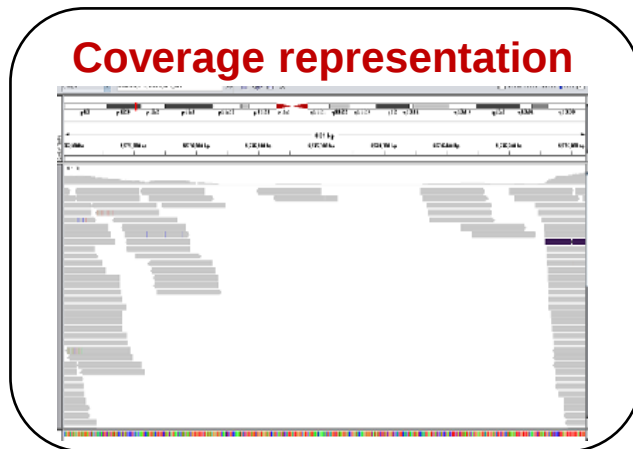
Variant Calling

```
GCTATGCATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACT
CTATGCATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTG
TATGCATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGT
ATGCATTGGCATGTCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTT
TGCATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTTA
GCATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTTAG
CATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTTAGC
ATTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTTAGCC
TTGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTTAGCCA
TGGCATGGCATGCTAGCTACGGGATGCTGATCGATTTTCGAAACTGACTGTTAGCCAT
```

reference GCTATGCATTGGCATGGCATGCTAGCTACAGGATGCTGATCGATTTTCGAAACTGACTGTTAGCCAT

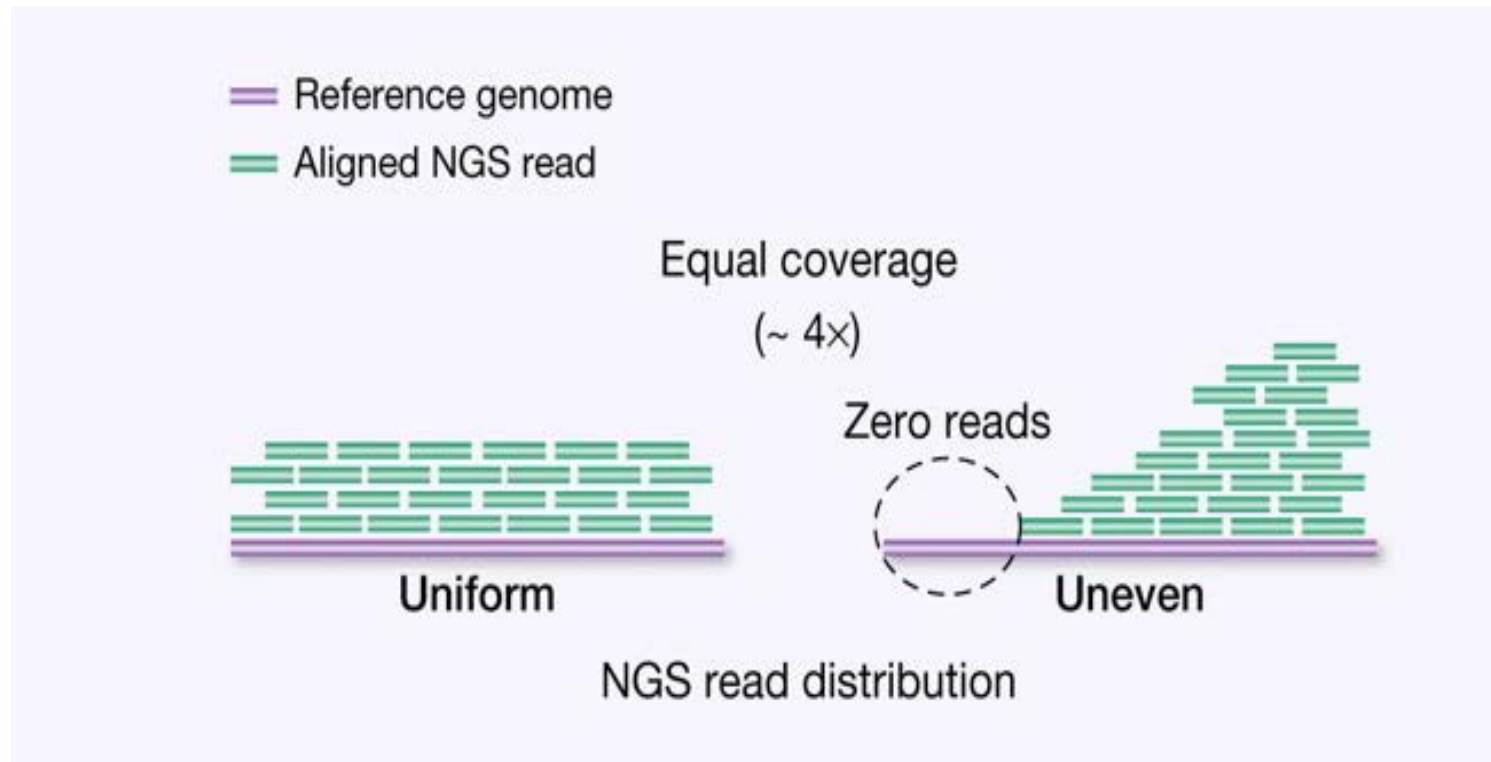
- NGS engine R-GenDx,
- Type Stream,
- Conexio Assign NGS,
- Omixon

Megabases-Gigabases



Περιορισμοί:

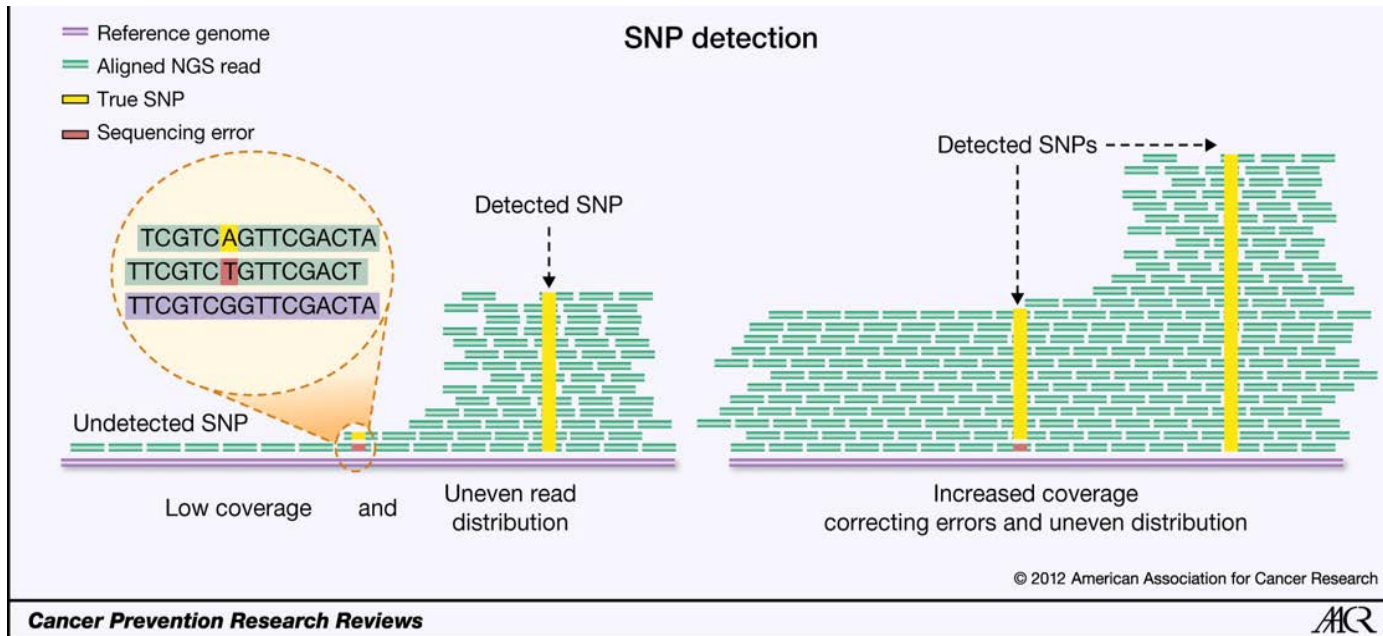
- ορισμένες περιοχές του ανθρωπίνου γονιδιώματος είναι πολύ δύσκολες ή αδύνατο να αλληλουχισθούν:
 - α) παρουσία **ψευδογονιδίων** (παρόμοια αλληλουχία με αυτή των γονιδίων)
 - β) ανεπαρκής ή χαμηλή ενίσχυση σε περιοχές **πλούσιες σε GC**, **ομοπολυμερείς** / **επαναλαμβανόμενες**
 - γ) περιοχές με υψηλότερα επίπεδα **ποικιλομορφίας** (Structural Variations, SVs) (απαιτείται κάλυψη ~ 20x-30x) (δυσκολίες στην ευθυγράμμιση)



Δυσκολία στην ευθυγράμμιση / συναρμολόγηση των σύντομων αναγνώσεων

Απαιτείται επαρκής αριθμός αλληλοεπικαλυπτόμενων αναγνώσεων
(“κάλυψη” ή “βάθος” αλληλούχισης)

Error rates & «Κάλυψη» ή «Βάθος» αλληλούχισης



- όλα τα βήματα ενίσχυσης εισάγουν **σφάλματα βάσεων** (0,01-16/100b)
- αδυναμία ανίχνευσης μεταλλάξεων χαμηλής συχνότητας
- η υψηλή κάλυψη αυξάνει την **δυνατότητα διάκρισης** μεταξύ **λάθους** & **ποικιλομορφίας**
- η ύπαρξη ατελών αλληλουχιών αναφοράς αυξάνει την πιθανότητα απόκλισης

Βάθος αλληλούχισης, κατανομή, μήκος των reads, ποιότητα της ακολουθίας:
διασφαλίζουν την ακρίβεια των αποτελεσμάτων

Κάθε τεχνολογία περιλαμβάνει πλατφόρμες με:

- ✓ μεταβαλλόμενη απόδοση και
- ✓ μήκη ανάγνωσης

Διαγνωστικής και Δημόσιας Υγείας Εργαστήρια

- ο **MiSeq της Illumina**

(εμπορικά διαθέσιμη)



- ο **S5 της Ion Torrent**

(εμπορικά διαθέσιμη)



(2^{ης} γενιάς) Illumina «MiSeq» vs Ion Torrent «S5»

Platform	Instrument	Reads per run	Avg Read length (pb)	Read Type	Error Type	Error Rate (%)	Data Generated per run (Gb)	Year
First Generation								
ABI Sanger	3730xl	96	400 – 900*	SE	NA	0.3	0.00069 to 0.0021	2002
Second Generation								
454	GS20	200	100	SE, PE	indel	1	0.02	2005
454	GS FLX	400	250	SE, PE	indel	1	0.1	2007
454	GS FLX Titanium	1M	450	SE, PE	indel	1	0.45	2009
454	GS FLX Titanium+	1M	700	SE, PE	indel	1	0.7	2011
454	GS Junior	100	400	SE, PE	indel	1	0.04	2010
454	GS Junior+	100	700	SE, PE	indel	1	0.07	2014
Illumina	MiniSeq	25M (maximum)	150	SE, PE	mismatch	1	7.5 (maximum)	2013
Illumina	MiSeq	25M (maximum)	300	SE, PE	mismatch	0.1	15 (maximum)	2011
Illumina	NextSeq	400M (maximum)	150	SE, PE	mismatch	1	120 (maximum)	2014
Illumina	HiSeq	5B (maximum)	150	SE, PE	mismatch	0.1	1.5Tb (maximum)	2012
Illumina	HiSeq X	6B (maximum)	150	SE, PE	mismatch	0.1	1.8Tb (maximum)	2014
SOLiD	5500 W	3B	75	SE	mismatch	~0.1	160	2011
SOLiD	5500xl W	6B	75	SE	mismatch	~0.1	320	2013
Ion Torrent	PGM 314 chip v2	400.000-550.000	400	SE	indel	1	0.06 to 0.1	2011
Ion Torrent	PGM 316 chip v2	2M - 3M	200	SE	indel	1	0.6 to 1	2011
Ion Torrent	PGM 318 chip v2	4M - 5.5M	400	SE	indel	1	1.2 to 2	2013
Ion Torrent	Ion Proton	60M - 80M	200	SE	indel	1	10	2012
Ion Torrent	Ion S5/S5XL 520	3M - 5M	400	SE	indel	1	1.2 to 2	2015
Ion Torrent	Ion S5/S5XL 530	15M-20M	400	SE	indel	1	3 to 5	2015
Ion Torrent	Ion S5/S5XL 540	60M - 80M	400	SE	indel	1	NA	2015
Third Generation								
PacBio	RS C1	432	1300	SE	indel	15	0.54	2011
PacBio	RS C2	432	2500	SE	indel	15	0.5 to 1	2012
PacBio	RS C2 XL	432	4300	SE	indel	15	0.5 to 1	2012
PacBio	RS II C2 XL	564	4600	SE	indel	15	0.5 to 1	2013
PacBio	RS II P5 C3	528	8500	SE	indel	13	0.5 to 1	2014
PacBio	RS II P6 C4	660	13500	SE	indel	12	0.5 to 1	2014
PacBio	Sequel	350	10000	SE	NA	NA	7	2016
Oxford Nanopore	MinION Mk	100	9545	1D,2D	indel/mismatch	12	1.5	2015
Oxford Nanopore	PromethION	NA	9846	1D,2D	NA	NA	2Tb to 4Tb	2016

3^{ης} γενιάς αλληλούχιση (Third Generation Sequencing, TGS)

ανάλυση μονόκλωνων DNA μορίων σε πραγματικό χρόνο (Single-Molecule Real Time, SMRT)

χωρίς προηγηθείσα κλωνική ενίσχυση (**Kb-Mb**)

α) Heliscope/Helicos: Single Sequencing Sequencing (SMS) (2010)

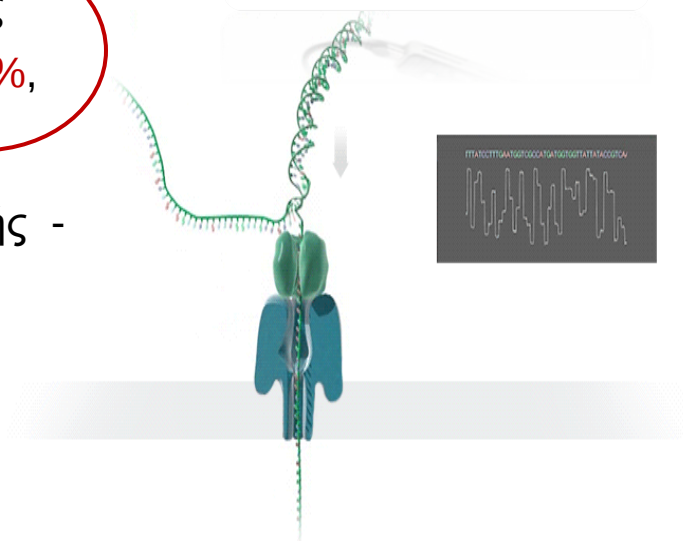
- αργή, ακριβή σε κόστος, σύντομα reads (~ 32 bp)
- δεν αποδείχθηκε βιώσιμη

β) Pacific Bioscience: Single Molecule Real-time (SMRT) (2011)

- απόδοση (260 Mb/διαδρομή), υψηλά ποσοστά σφαλμάτων (**13%**)
- μήκος reads ~ 40 Kb
- αλληλούχιση μεγάλων γενετικών περιοχών

γ) Oxford Nanopore τεχνολογία MinION (2014)

- φορητή συσκευή (USB stick μνήμης)
- μέτρηση της μεταβολής της ηλεκτρικής αγωγιμότητας
- μήκος reads ~50 Kb, χαμηλό κόστος, σφάλματα **~12%**, απόδοση 400 Mb/48h
- **PromethION“ (2016):** 48 ανεξάρτητες κυψελίδες ροής -
 - ισοδύναμο με 48 MinIONs
 - λειτουργεί στα **500bp/sec**
 - είναι ικανό να παρόξει **~2-4 Tb / 2 ημέρες**
 - μήκος reads **200Kb**



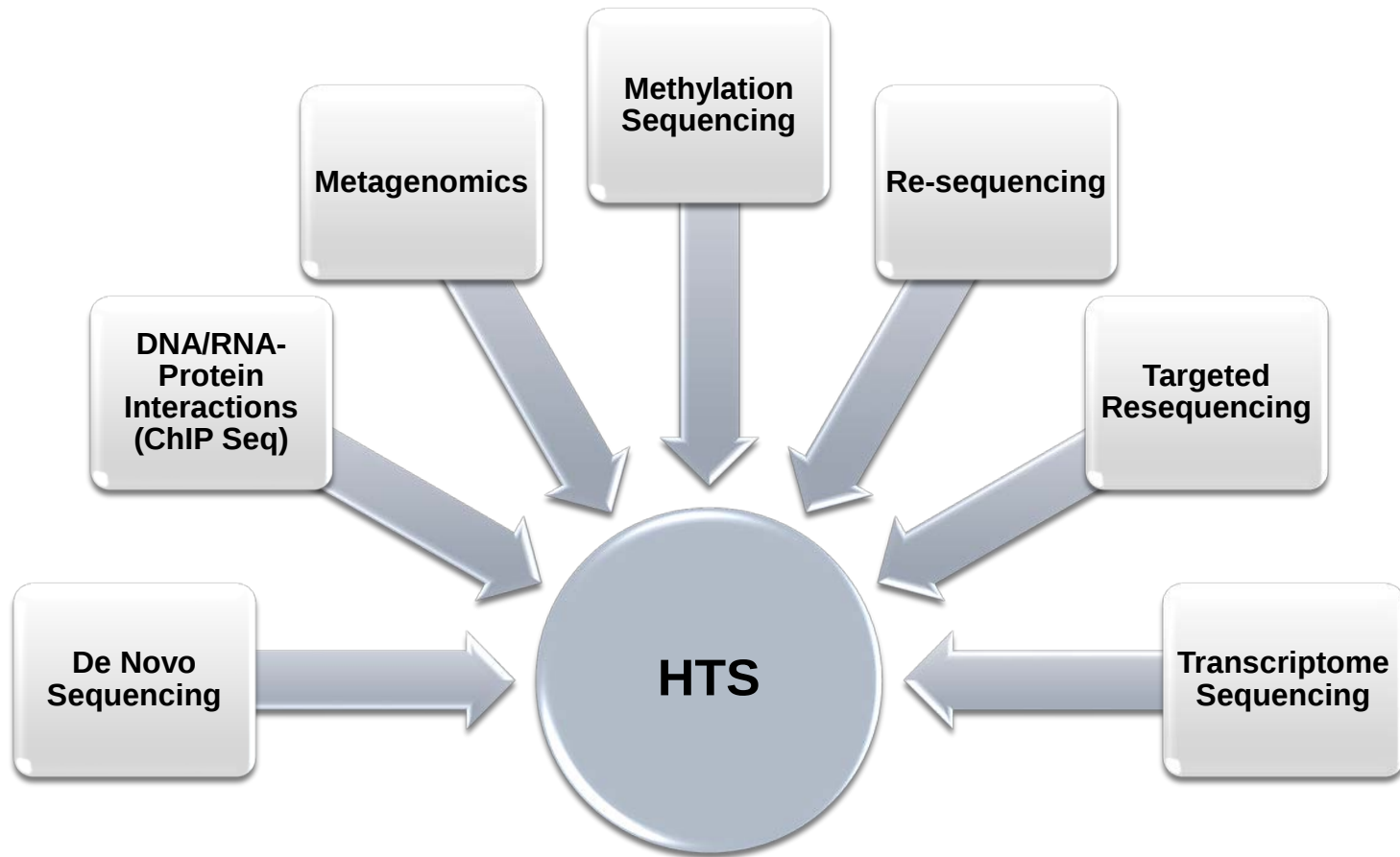
❑ αλληλούχιση 2^{ης} γενιάς:

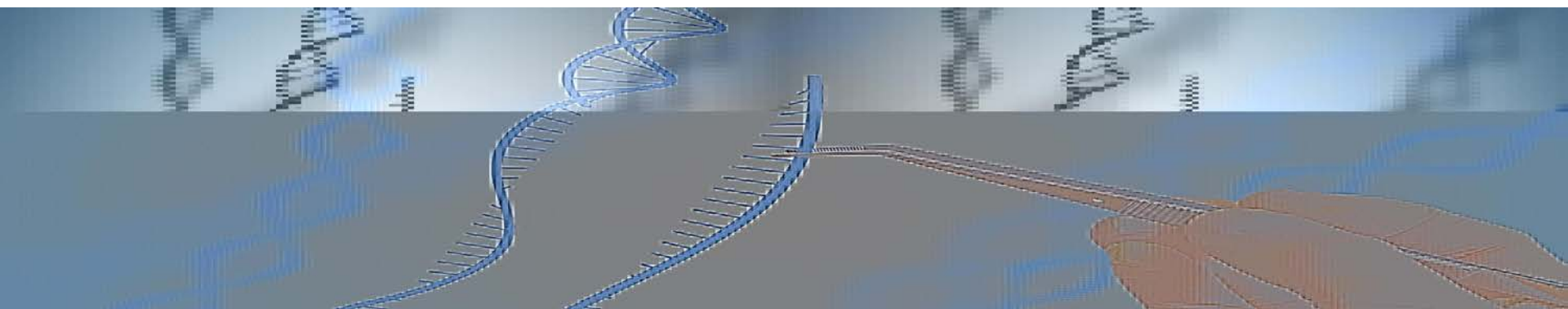
- ✓ μείωση κόστους αλληλούχισης
- ✓ μείωση χρόνου προετοιμασίας βιβλιοθήκης
- ✓ μείωση κόστους μεταφοράς και ανάλυσης δεδομένων

❑ αλληλούχιση 3^{ης} γενιάς:

Πρότυπο Ιατρικό Διαγνωστικό Εργαλείο στο εγγύς μέλλον

- ✓ θα επιτρέψει την πληρέστερη, χωρίς κενά συναρμολόγηση ανθρ.γονιδιωμάτων
- ✓ θα ενισχύσει περαιτέρω την έρευνα στην ανθρώπινη γενετική
- ✓ εξατομικευμένη ιατρική
- ✓ φορητότητα του MinION : αλληλούχιση στις αναπτυσσόμενες χώρες





The Answer is in the Sequence

Σας ευχαριστώ