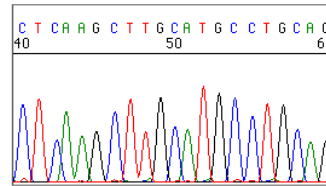
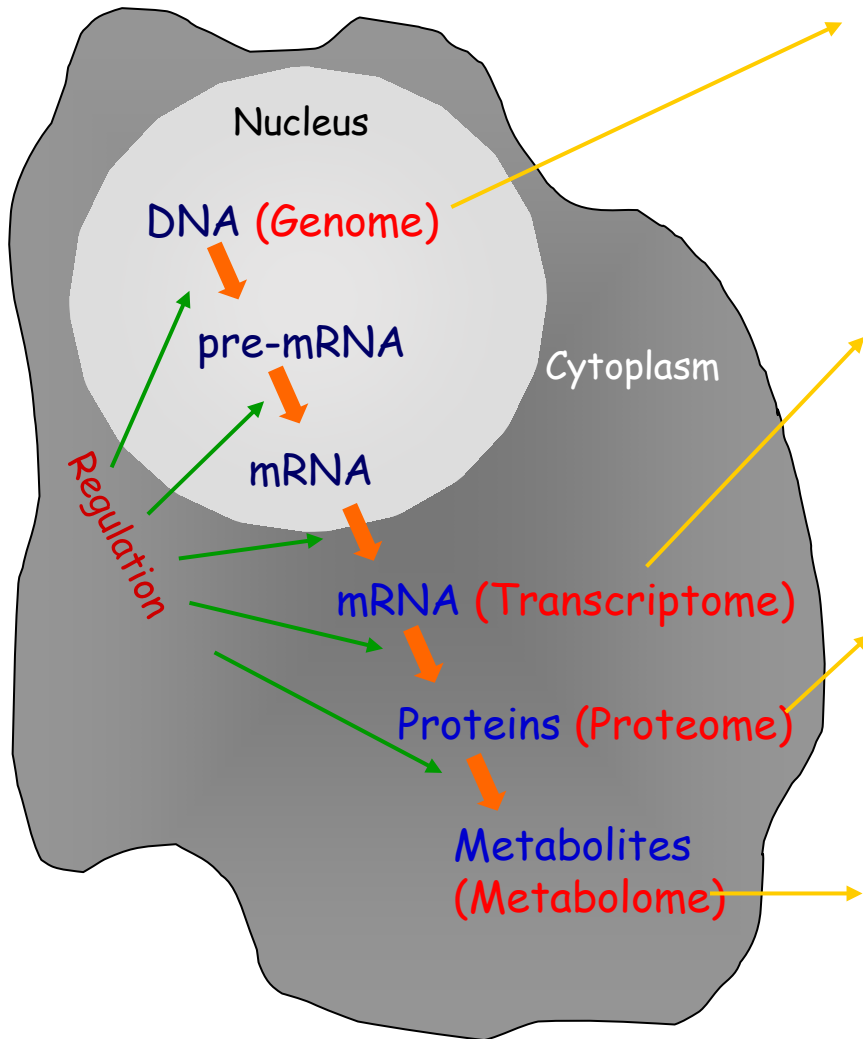


Από το
γονίδιο στο
γονιδίωμα

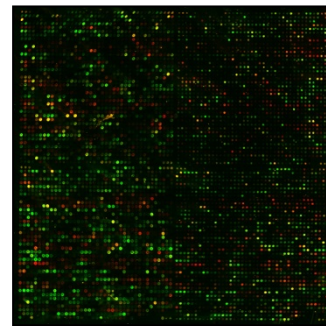


ΓΟΝΙΔΙΑΚΗ ΕΚΦΡΑΣΗ



- **Genome mapping**
- **Genome sequencing**
- **Genome annotations**

Structural genomics

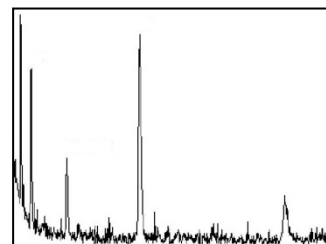


- **DNA arrays and chips**
- **RNA sequencing**
- (semi) qRT-PCR
- Northern blot + hybrid.
- Transcriptional fusions

Functional genomics



- **2D electrophoresis**
- **Gel-free methods**
- Mass spectrometry
- Protein sequencing
- Translational fusional
- Immunodetection
- Enzyme activities



- **Chromatography**
- **Mass spectrometry**
- **NMR**

- **Γονιδίωμα:** περιλαμβάνει όλη την πληροφορία στο σύνολο των χρωμοσωμάτων που συναντώνται στον πυρήνα (ή και έξω από αυτόν) κάθε κυττάρου ενός οργανισμού

-
- **Η Γονιδιωματική Επιστήμη (Genomics)** έχει σκοπό της τη μελέτη και την ανάλυση ολόκληρων γονιδιωμάτων
 - Ο όρος χρησιμοποιήθηκε πρώτη φορά το 1986.
 - Εφαρμογή εργαστηριακών μεθόδων χαρτογράφησης, εύρεσης αλληλουχίας, σύγκρισης αποτελεσμάτων και εξαγωγής συμπερασμάτων

Plant Genome

Diversity, Conservation
and Manipulation

Editors
B.K. Roy
B.R. Chaudhary
R.P. Sinha



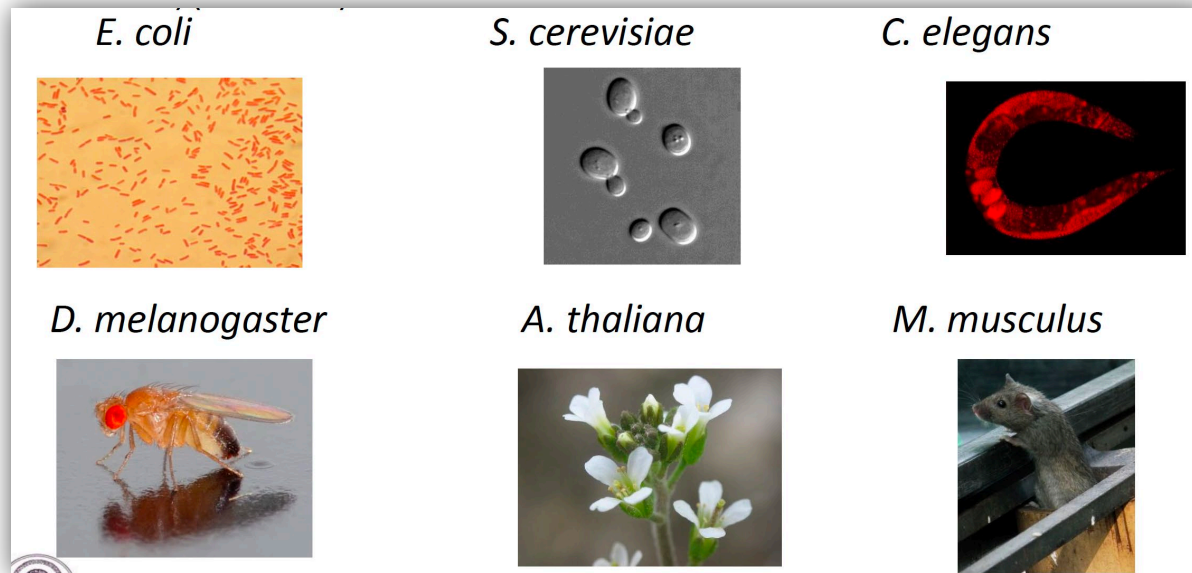
Narosa

- ❖ **Δομική Γονιδιωματική (Structural Genomics)** - αναλύει τη δομή του γονιδιώματος (γενετική χαρτογράφηση, φυσική χαρτογράφηση, αλληλούχηση). Μπορεί να περιλαμβάνει την μελέτη των αλληλεπιδράσεων μεταξύ γονιδίων (genetic interactions – epistasis) αλλά και γονιδίων περιβάλλοντος (epigenomics) Αναπτύσσεται ραγδαία

- ❖ **Συγκριτική Γονιδιωματική (Comparative Genomics)** - Εφαρμογή της γνώσης από ένα είδος σε άλλους οργανισμούς (ακόμα και απομακρυσμένους εξελικτικά)
 - Οι πληροφορίες συλλέγονται από οργανισμούς μοντέλα και έχουν προεκτάσεις σε γεωργικά και ιατρικά προβλήματα

 - Ανάλυση της σχέσης γενότυπου-φαινότυπου

Οργανισμοί Μοντέλα



- Έχουν **πλήθος κοινών γενετικών μηχανισμών** και μονοπατιών μεταξύ τους και με τον άνθρωπο.
- Είναι δυνατά **πειράματα ελεγχόμενων διασταυρώσεων και κατευθυνόμενης τροποποίησης** γενετικού υλικού
- Δεν επιλέχτηκαν τυχαία – Βρίσκονται σε **διαφορετικά σκαλοπάτια** της εξελικτικής αλυσίδας.
- Υπάρχουν πολλά βιολογικά δεδομένα και μελετώνται για πολλά χρόνια

Τα πρώτα προγράμματα αλληλούχησης

Οργανισμός	# Γονιδίων	% γονιδίων με γνωστή λειτουργία	Ημερομηνία ολοκλήρωσης
<i>E. coli</i>	4.288	60	1997
<i>S. cerevisiae</i>	6.600	40	1996
<i>C. elegans</i>	20.000	40	1998
<i>Drosophila</i>	15.000	25	1999
<i>Arabidopsis</i>	25.000	40	2000
Ποντίκι	~30.000	20	2002
Άνθρωπος	~30.000	20	2001/4

Τα κύρια συμπεράσματα που προκύπτουν από ολοκληρωμένα ή ημι-ολοκληρωμένα προγράμματα αλληλούχησης γονιδιωμάτων διαφόρων ειδών, όπως:

- αριθμός και ο τύπος των γονιδίων,
- το ποσοστό των επαναλαμβανόμενων ακολουθιών,
- η οργάνωση και η δομή των γονιδιωμάτων,
- η εξέλιξή τους, και
- οι μελλοντικές κατευθύνσεις της έρευνας

Ανάλυση γονιδιώματων

Οι ερευνητές στην προσπάθειά τους να μελετήσουν το γονιδίωμα ενός οργανισμού ενδιαφέρονται για δύο κύρια στοιχεία του:

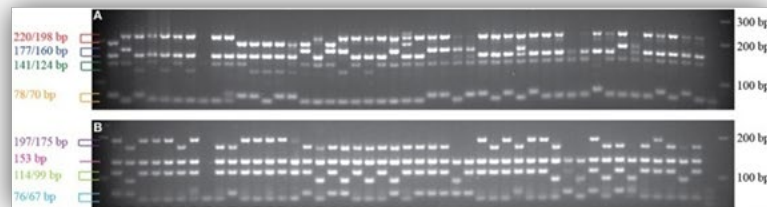
- Την **εύρεση της αλληλουχίας** κατά μήκος όλων των χρωμοσωμάτων ενός απλοειδούς γονιδιώματος.
- Την **αποκάλυψη της ποικιλομορφίας** ή αλλιώς της ύπαρξης διαφορετικών αλληλομόρφων που συναντώνται μεταξύ των διαφορετικών ατόμων του ίδιου είδους ή και στις δύο σειρές ομόλογων χρωμοσωμάτων (αν αναφερόμαστε σε διπλοειδή οργανισμό).

```
ACGAGTCGGTAGCTGCCCTCTGACTGCATCGAATTG  
CTCCCTACTACGTGCTATATGCGCTTACGATCGTAC  
GAAGATTTATAGAATGCTGCTACTGCTCCCTTATTTCG  
ATAACTAGCTCGATTATAGCTACGATG
```



Μοριακοί δείκτες (Molecular markers)

- **Γενετικοί δείκτες** (DNA, ή προερχόμενοι από RNA (ESTs) συσχέτιση με συγκεκριμένα γονίδια, γενετικούς τόπους, χρωμοσώματα, πολυμορφισμούς, έκφραση γονιδίων. Εφαρμογές στην Βελτίωση Φυτών σύνδεση με φαινοτυπικά γνωρίσματα και γενικότερα στη γενετική ανάλυση - ταυτοποίηση φυτικών και όχι μόνο οργανισμών)
- **Βιοχημικοί δείκτες** (πρωτεΐνες, σηματοδοτικά μόρια, μόρια ενέργειας, ιόντα, κλπ.)
- Ζητούμενο η υψηλή εξειδίκευση, η εύκολη ανίχνευση και η ταυτοποίηση
- Σημαντικότερο ζητούμενο η σωστή ποσοτικοποίηση και η επαναληψιμότητα



Προβλήματα στην αλληλούχηση

Τα προβλήματα που μπορούν να προκύψουν στην αλληλούχηση είναι τα εξής:

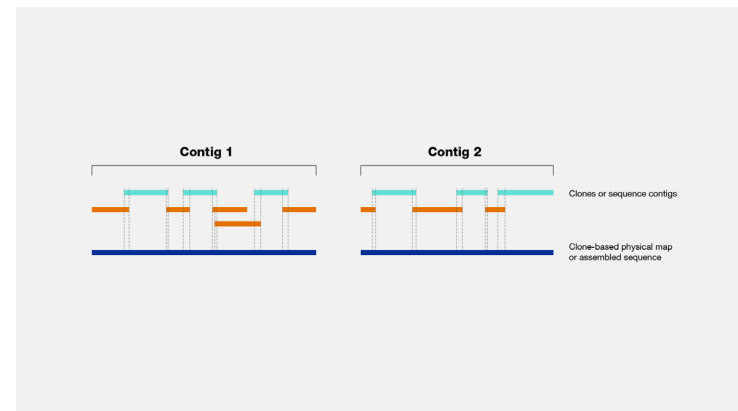
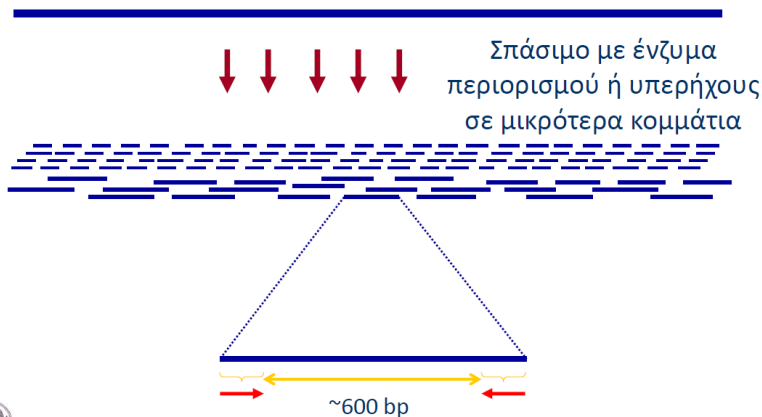
- ✓ Τα λάθη κατά την αλληλούχηση
- ✓ Ο πολυμορφισμός των γονιδιωμάτων (ετεροζυγωτία) (διπλοειδείς
- ✓ Οι επαναλαμβανόμενες αλληλουχίες που εντοπίζονται με δυσκολία
 - *GT επαναλήψεις διάσπαρτες στο γονιδίωμα μεγέθους >600 βάσεις.*
- ✓ Τα τμήματα DNA που δεν κλωνοποιούνται και συνεπώς δεν μπορούν και να αλληλουχηθούν
 - *Ιδιαίτερα οι ετεροχρωματινικές περιοχές του DNA δεν κλωνοποιούνται. Αν και η ετεροχρωματίνη περιέχει πολύ λίγα γονίδια αποτελεί περίπου το 30% του συνολικού DNA.*

Λύση:

Τα χρωμοσώματα σπάζουν σε πολύ μικρά και αλληλοεπικαλυπτόμενα κομμάτια και κλωνοποιούνται σε βιβλιοθήκες.

Η αλληλούχηση μπορεί να επιτευχθεί κατά μήκος όλων σχεδόν των κομματιών. Αφού υπάρχει αλληλοεπικάλυψη, είναι δυνατή η επανασυναρμολόγηση των επιμέρους κομματιών στο συνολικό χρωμόσωμα

Γονιδίωμα

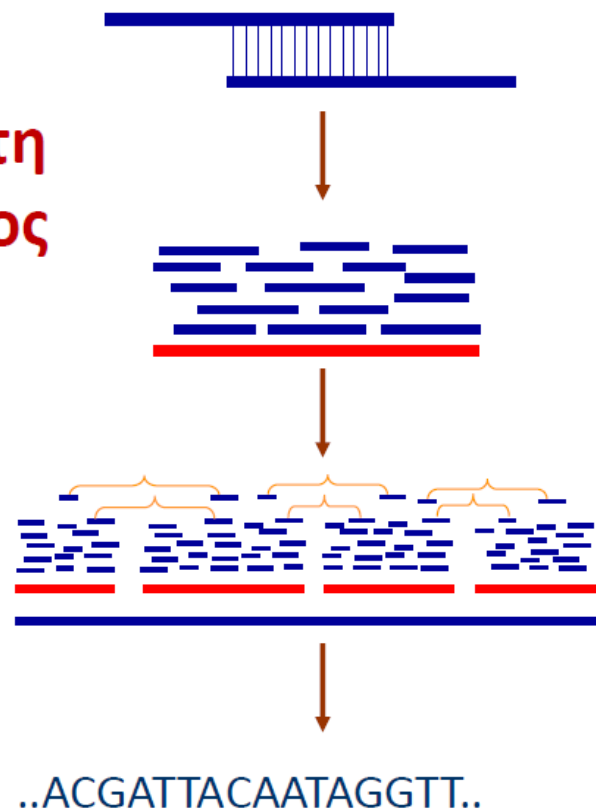


Υπολογιστικά εργαλεία

Υπολογιστικά προγράμματα για τη συναρμολόγηση του γονιδιώματος

PHRED, PHRAP, LUCY,
TGICL, CAP3, GAP4

<http://seqanswers.com/wiki/Software/list>,
<http://www.jcvi.org/cms/research/software/>



Αλληλούχηση γονιδιώματος

Proc. Natl. Acad. Sci. USA
Vol. 74, No. 12, pp. 5463–5467, December 1977
Biochemistry

DNA sequencing with chain-terminating inhibitors

(DNA polymerase/nucleotide sequences/bacteriophage ϕ X174)

F. SANGER, S. NICKLEN, AND A. R. COULSON

Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

Contributed by F. Sanger, October 3, 1977

ABSTRACT A new method for determining nucleotide sequences in DNA is described. It is similar to the “plus and minus” method [Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* 94, 441–448] but makes use of the 2',3'-dideoxy and arabinonucleoside analogues of the normal deoxynucleoside triphosphates, which act as specific chain-terminating inhibitors of DNA polymerase. The technique has been applied to the DNA of bacteriophage ϕ X174 and is more rapid and more accurate than either the plus or the minus method.





Charles Darwin published "On the Origin of Species by means of Natural Selection,"
1859



Gregory Mendel introduced the fundamental laws of inheritance
1865



Chromosomes and cancer relationship has been proposed by Boveri
1902



James Watson and Francis Crick described the structure of DNA
1953



Sanger sequencing method was developed
1977

Invention of "polymerase chain reaction" by Kerry Mullis
1985



"Human Genome Project" was launched
1990

The first draft of "Human Genome Project" was reported
2001

"Human Genome Project" was officially completed
2003

Applied biosystems, Illumina, Roche Company, Pacific Biosciences, Oxford Technologies Nanopore, Helicos Biosciences, and Solexa launched 2nd and 3rd generation sequencing platforms
2004-

1665
"Cell" was described by Robert Hooke

1888
"Chromosome" was described by Waldeyer



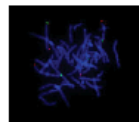
1910
Thomas Hunt Morgan showed that genes are located on chromosomes

1956
Levan and Tijo reported the human chromosome number was 46

1959
Trisomy 21 was described in Down syndrome by Lejeune



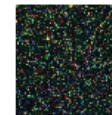
1982
Fluorescence in situ hybridization (FISH) was developed




1980
Maxam-Gilbert sequencing method was developed

1992
Comparative genomic hybridization (CGH) was developed

2000
Massively parallel sequencing (MPS) was developed by Lynx Therapeutics





1869 – Discovery of DNA

1909 – Chemical characterisation

1953 – Structure of DNA solved

1977 – Sanger sequencing invented

– First genome sequenced – Φ X174 (5 kb)

1986 – First automated sequencing machine

1990 – Human Genome Project started

1992 – First “sequencing factory” at TIGR

1995 – First bacterial genome – *H. influenzae* (1.8 Mb)

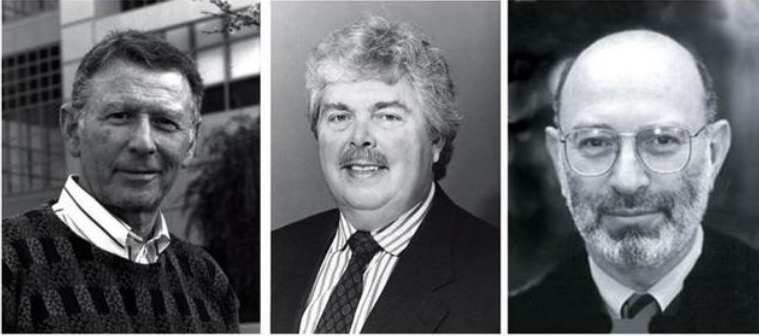
1998 – First animal genome – *C. elegans* (97 Mb)

2003 – Completion of Human Genome Project (3 Gb)

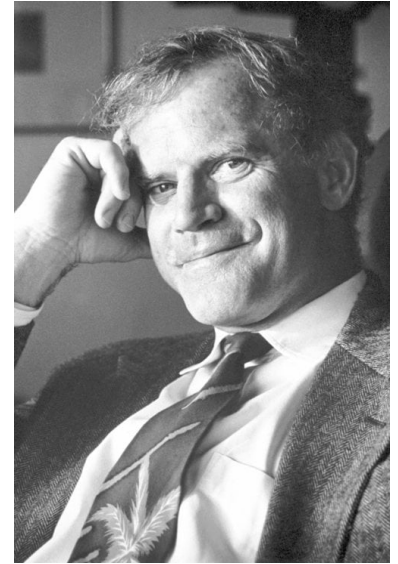
– 13 years, \$2.7 bn

2005 – First “next-generation” sequencing instrument - 454 Life Sciences

2013 – >10,000 genome sequences in NCBI database



P. Burg, H. Boyer, S. Cohen



K. Mullis

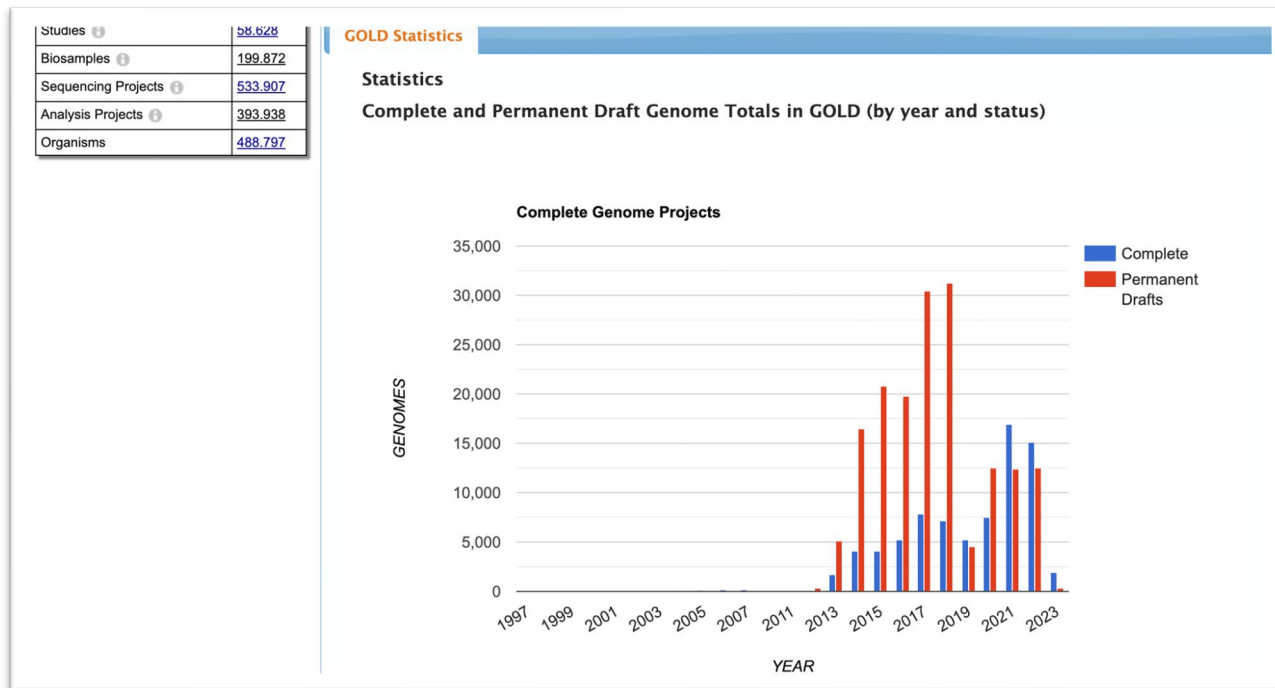
1977 bacteriophage ϕ X174 (5386bp, 11 genes)

- 1981 mitochondrial genome (16,568bp; 13 prots; 2 rRNAs; 22 tRNAs)
- 1986 chloroplast genome (120,000-200,000bp)
- 1995 *Haemophilus influenzae* (1.8Mb)

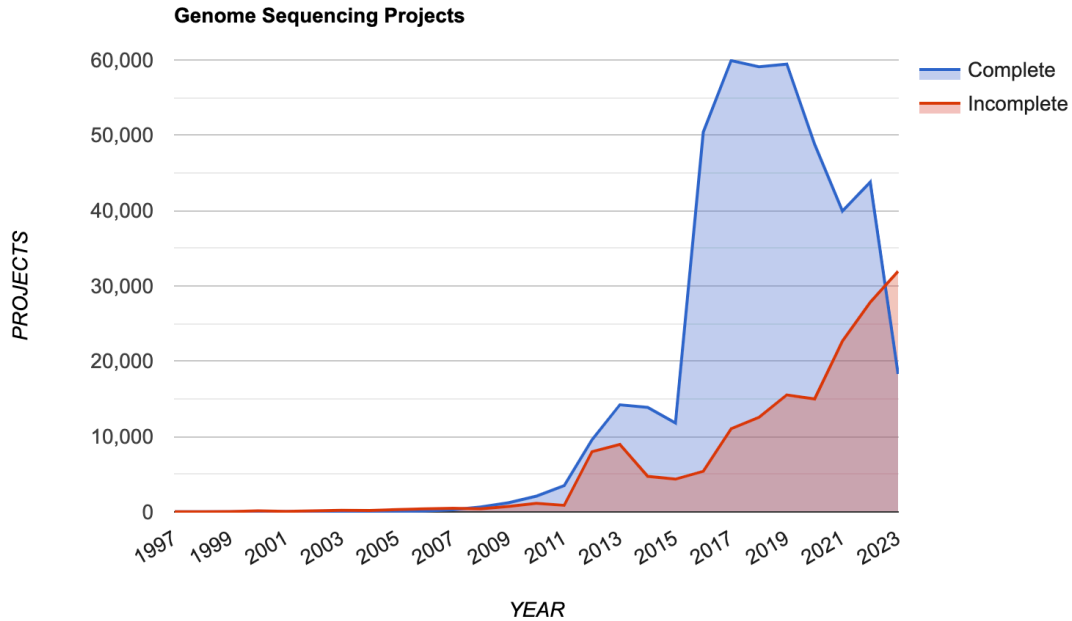
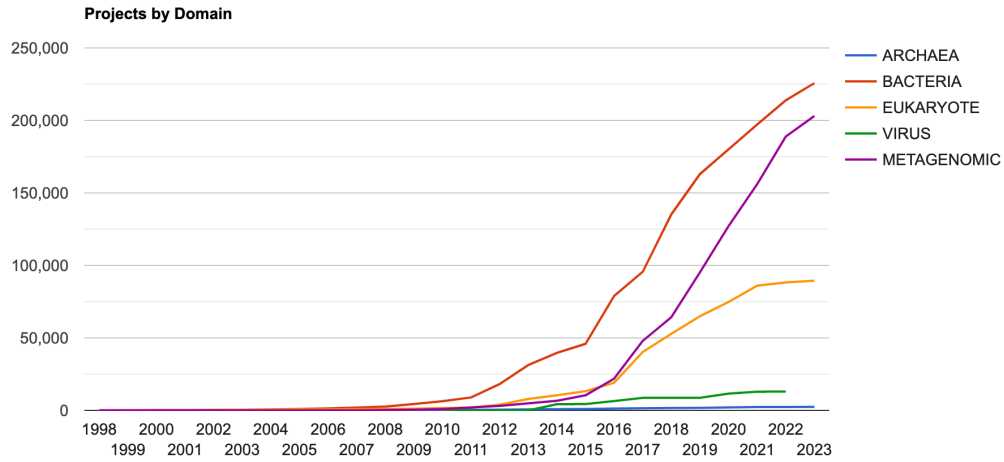
- 1996 *Saccharomyces* whole genome (12.1Mb; over 600 people 100 laboratories)
- 1997 *E. coli* (4.6Mb; 4200 proteins)
- 1998 *Caenorhabditis elegans* (97 Mb; 19,000 genes)
- 2000 *Arabidopsis thaliana* (115Mb, 25-30,000 genes)
- 2001 mouse (1 year!)



GOLD (Genomes online database) είναι μια ηλεκτρονική πηγή που παρέχει ολοκληρωμένη πρόσβαση σε πληροφορίες που αφορούν γονιδιωματικά και μεταγονιδιωματικά προγράμματα αλληλούχησης, καθώς και των δεδομένων που προκύπτουν από αυτά

<https://gold.jgi.doe.gov/index>



Project Totals in GOLD (by year and Domain Group)



- 
-
- ✓ Σήμερα σε 31 χώρες παγκοσμίως παράγονται δεδομένα από γονιδιωματικές έρευνες
 - ✓ Στον παρακάτω σύνδεσμο φαίνονται οι χώρες αυτές, καθώς και ο αριθμός των προγραμμάτων σε καθεμία από αυτές:
http://genomesonline.org/cgi-bin/GOLD/index.cgi?page_requested=GenomeMap
 - ✓ Η.Π.Α: Η χώρα με τα περισσότερα γονιδιωματικά προγράμματα (κατέχει το προβάδισμα με διαφορά από τις υπόλοιπες, > 3.000 προγράμματα)
 - ✓ Ακολουθούν το Η. Βασίλειο (~ 250 προγράμματα), η Ιαπωνία (~ 200 προγράμματα) και η Γαλλία (~ 200 προγράμματα)
 - ✓ Μεταξύ των σημαντικότερων κέντρων αλληλούχισης περιλαμβάνονται τα: JGI (Joint Genome Institute), JCVI (J. Craig Venter Institute), Broad Institute
http://genomesonline.org/cgi-bin/GOLD/index.cgi?page_requested=Statistics
- 

Major Sequencing Centers for Plant Genomes

Major Sequencing Centers for Plant Genomes

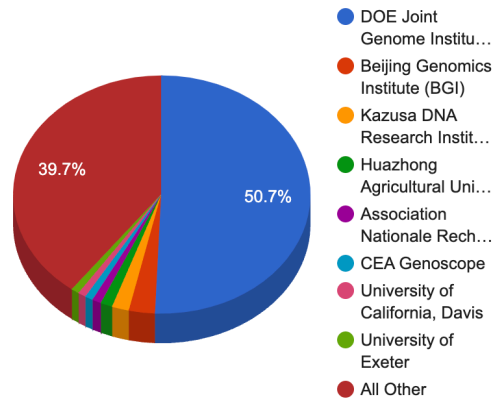


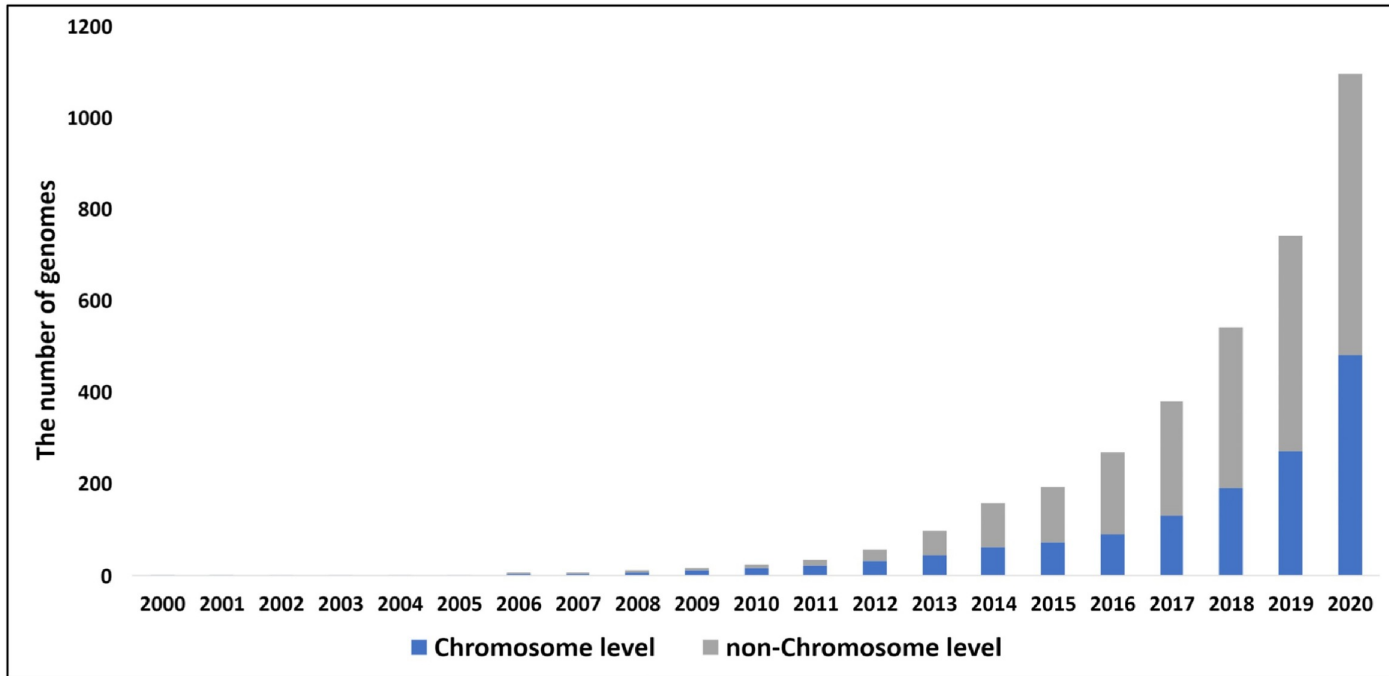
Fig. 3: Geographic distribution of the submitting institutions for 798 plant genome assemblies.

From: [Representation and participation across 20 years of plant genome sequencing](#)

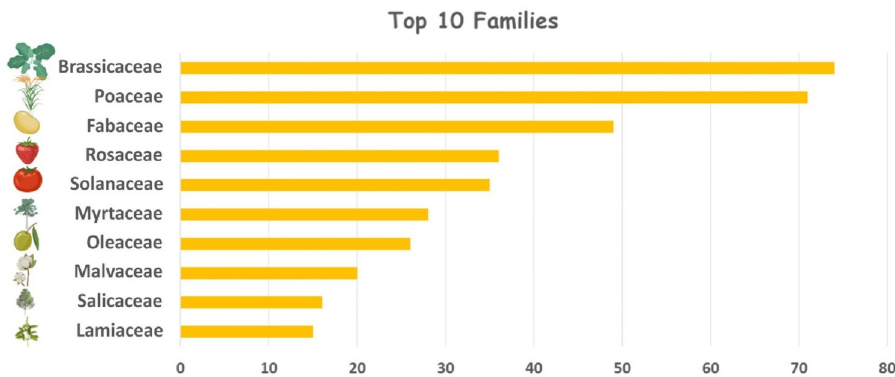


Circles are scaled by the number of genome assemblies produced in each nation and coloured by the relative proportion of domesticated, cultivated, feral, natural commodity, wild and wild relative species sequenced.

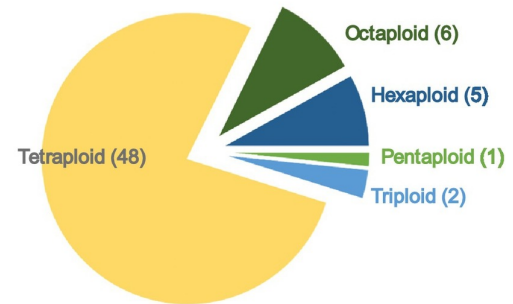
(A)



(B)



(C)



Trends in Plant Science

Οι πρώτες απόπειρες αλληλούχησης
μορίων χαρακτηρίστηκαν ως
"Αλληλούχηση 1ης γενιάς" - *1st Generation
Sequencing*,

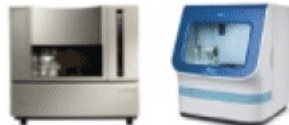
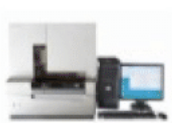
στην συνέχεια έκαναν την εμφάνισή τους
τεχνικές 2ης Γενιάς - *2nd Generation
Sequencing*

και τέλος 3ης ή «Επόμενης Γενιάς» - *Next
Generation Sequencing* οι οποίες σήμερα
χαρακτηρίζονται από τη δυνατότητα
μαζικής αλληλούχησης, υψηλής ταχύτητας
αλλά κυρίως από χαμηλό κόστος

First generation

Second generation (next generation sequencing)

Third generation



Sanger sequencing
Maxam and Gilbert
Sanger chain termination

Infer nucleotide identity using dNTPs,
then visualize with electrophoresis

500–1,000 bp fragments

454, Solexa,
Ion Torrent,
Illumina

High throughput from the
parallelization of sequencing reactions

~50–500 bp fragments

PacBio
Oxford Nanopore

Sequence native DNA in real time
with single-molecule resolution

Tens of kb fragments, on average

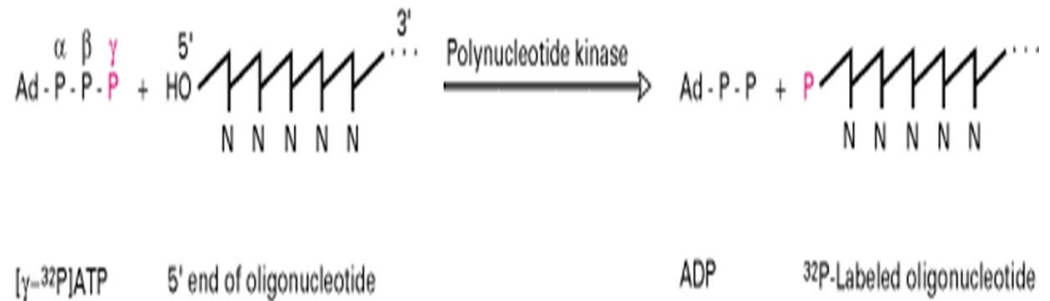
Short-read sequencing

Long-read sequencing

Αλληλούχηση χημικής διάσπασης Maxam- Gilbert

Η τεχνική βασίζεται στην χημική διάσπαση ενός επισημασμένου μορίου DNA σε συγκεκριμένες θέσεις όπου υπάρχει αδενίνη, γουανίνη, κυτοσίνη ή θυμίνη.

Μετά από αυτή τη μερική διάσπαση έχουμε ένα σύνολο ραδιενεργών θραυσμάτων, διαφορετικού μεγέθους. Έπειτα από ηλεκτροφόρηση σε τζελ πολυακρυλαμιδίου και αυτοραδιογραφία διαχωρίζονται αυτά τα μονόκλιωνα θραύσματα κατά μέγεθος.



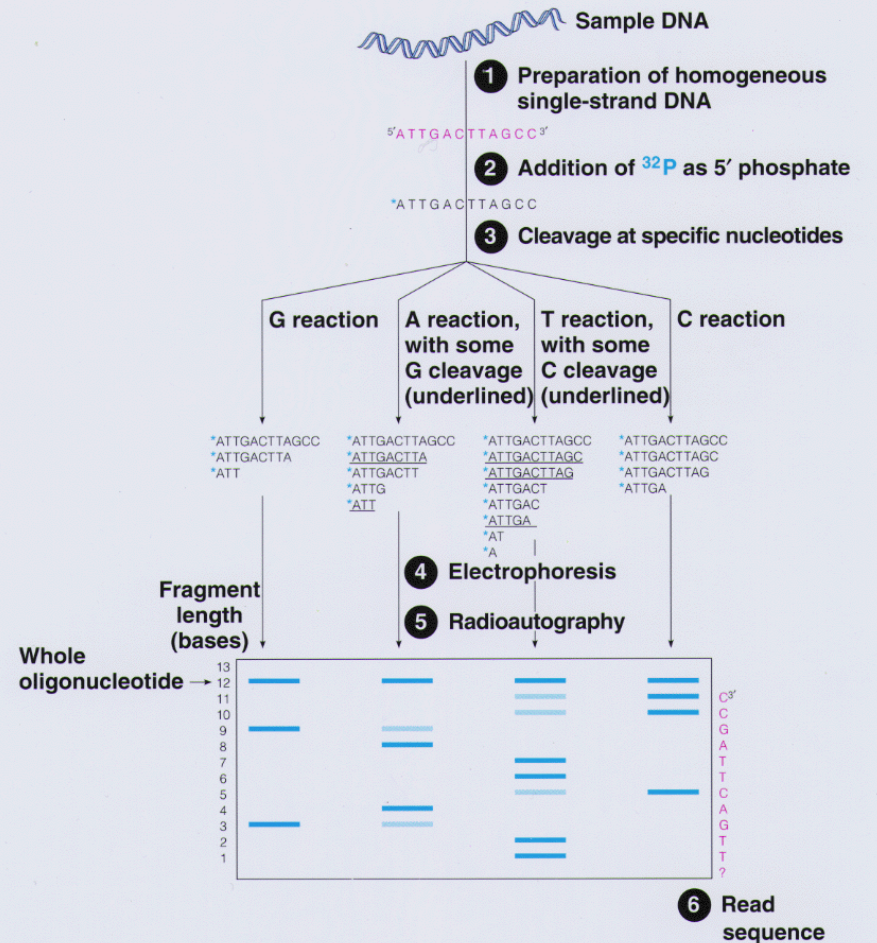
Η διάσπαση γίνεται με αντιδραστήρια τα οποία πρώτα εκτελούν την τροποποίηση της βάσης, έπειτα απομακρύνουν την βάση από το σάκχαρο της. Το σάκχαρο που είναι εκτεθειμένο τότε σπάει.

Τα θραύσματα από τις τέσσερις αντιδράσεις ηλεκτροφορούνται και έπειτα οπτικοποιούνται με έκθεση σε ακτίνες X για την ανάγνωση της αλληλουχίας.

Chemical breakdown

Purines		Pyrimidines	
DMSO + Piperidine		Hydrazine + Piperidine	
1	2	3	4
Alkaline	Acidic (formic acid)	high salt	low salt
• Cleaves specifically at G	• Cleaves at G or A	• Cleaves specifically at C	• Cleaves at C or T

Figure 4A.4 Sequencing an oligonucleotide by the Maxam-Gilbert method



Μέθοδος Sanger



Frederick Sanger

Η διδέοξυ μέθοδος ή μέθοδος τερματισμού της αλυσίδας (chain terminator ή αλλιώς μέθοδος Sanger).

Το κλειδί της μεθόδου είναι η χρήση διδέοξυ-νουκλεοτιδίων (ddNTPs) για τον τερματισμό της αλυσίδας του DNA.

Απαιτείται μία μονόκλωνη αλυσίδα DNA ως καλούπι (template), ένας εκκινήτης (primer), μία DNA πολυμεράση, ραδιοσημασμένα ή φθορίζοντα νουκλεοτίδια και τροποποιημένα νουκλεοτίδια που τερματίζουν την επιμήκυνση του DNA.

Το δείγμα του DNA διαιρείται σε 4 ξεχωριστές αντιδράσεις αλληλούχησης, που η καθμία περιέχει τα νουκλεοτίδια (dATP, dGTP, dCTP and dTTP), την πολυμεράση και ένα από τα 4 διδεοξυνουκλεοτίδια (ddATP, ddGTP, ddCTP, ddTTP), **τα οποία στερούνται της 3'-OH που απαιτείται για το σχηματισμό του φωσφοδιεστερικού δεσμού μεταξύ δύο νουκλεοτιδίων.**

(a)

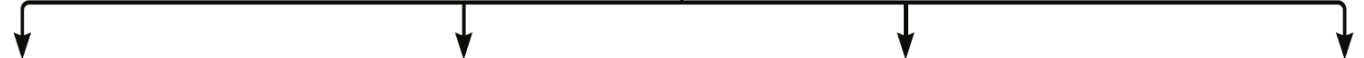
Δίκλωνο DNA
••CGATGTACGTCTAGG••
••GCTACATGCAGATCC••

Δημιουργία μονόκλωνης μήτρας

••GCTACATGCAGATCC••

Πραγματοποίηση τεσσάρων αντιδράσεων

Νουκλεοτίδιο



A

T

G

C

Μείγμα νουκλεοτιδίων

dATP
dTTP
ddTTP
dGTP
dCTP

dATP
ddATP
dTTP
dGTP
dCTP

dATP
dTTP
dGTP
dCTP
ddCTP

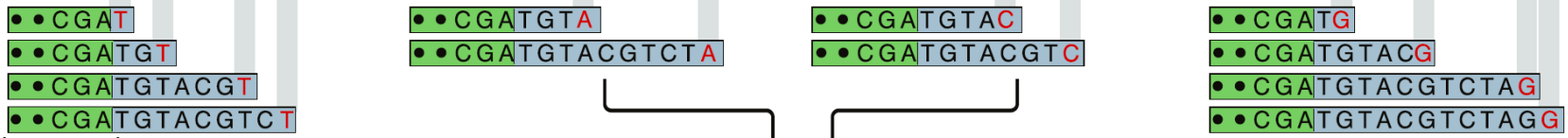
dATP
dTTP
dGTP
ddGTP
dCTP

Προσθήκη DNA πολυμεράσης



••GCTACATGCAGATCC•• ••GCTACATGCAGATCC•• ••GCTACATGCAGATCC•• ••GCTACATGCAGATCC••

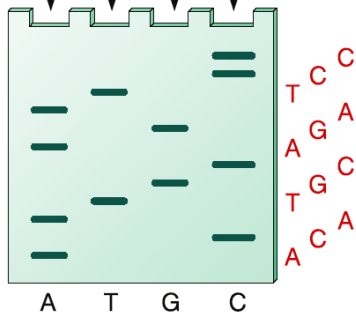
(β)



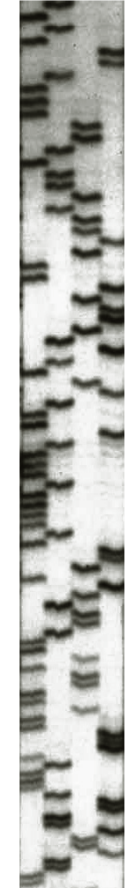
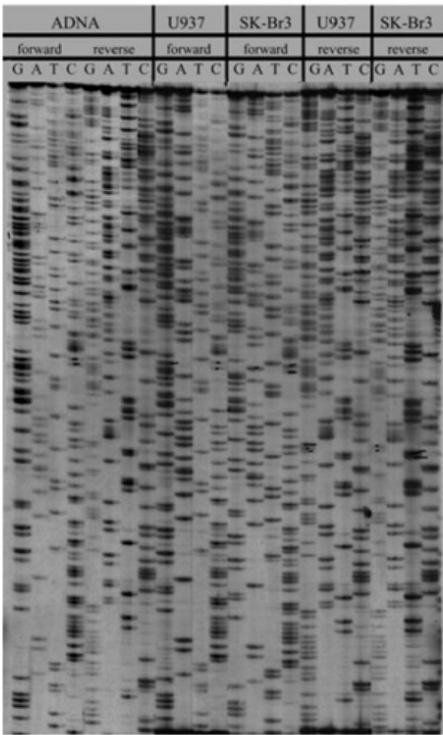
Εκκινητής

(γ)

Τα προϊόντα κάθε αντίδρασης αναλύονται σε μία στήλη ενός ηλεκτροφόρου πηκτώματος πολυακρυλαμίδης

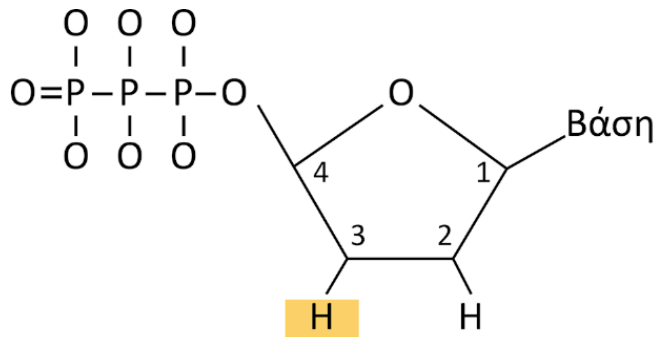


Αναγνωσμένη αλληλουχία
ACATGCAGATCC

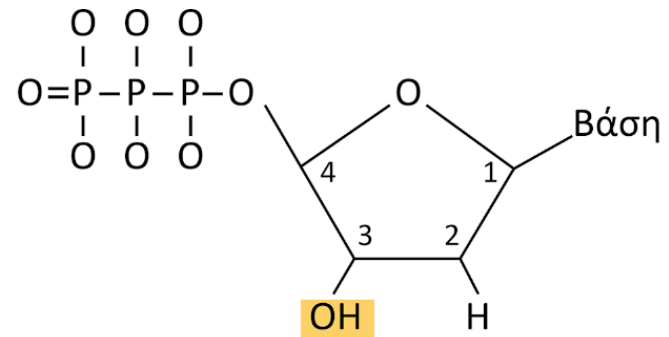


GATC

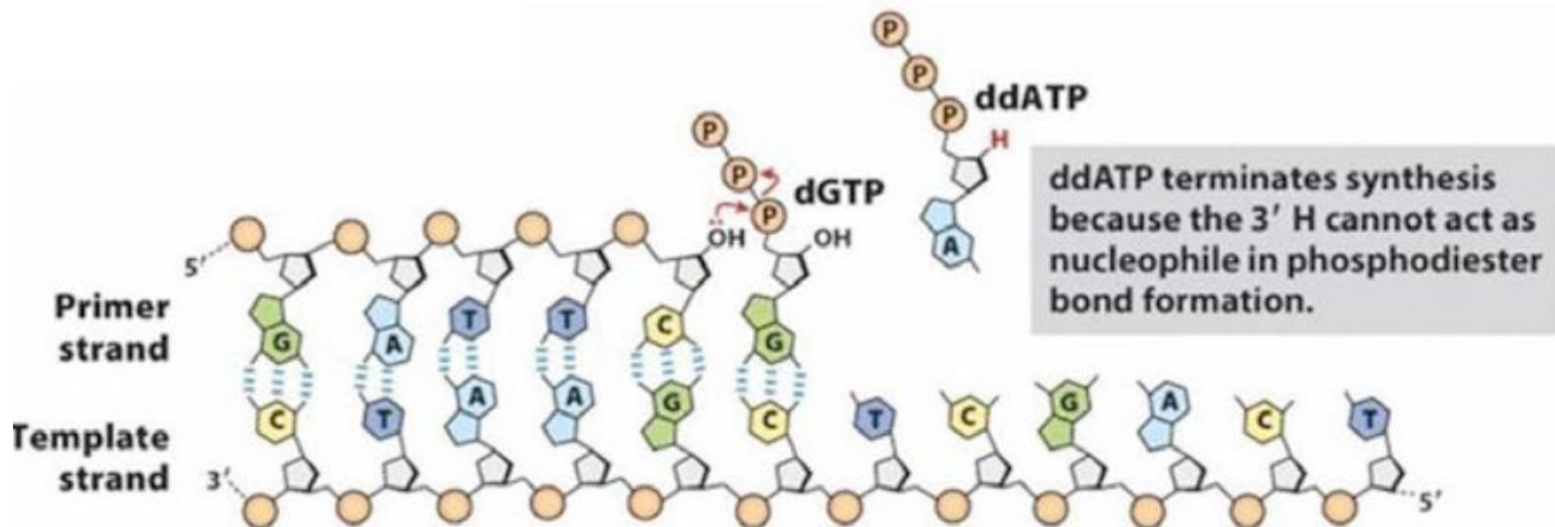
Τα τριφωσφορικά 2', 3'-διδεοξυνουκλεοτίδια (ddNTPs) φέρουν υδρογόνο αντί υδροξυλίωσης θέση 3' της δεοξυριβόζης το οποίο είναι απαραίτητο για να δημιουργηθεί ο φωσφοδιεστερικός δεσμός.



Διδεοξυνουκλεοτίδιο (ddNTP)



Δεοξυνουκλεοτίδιο (dNTP)



Template: 3' ——— CCGGTAGCAACT ——— 5'
 Primer: 5' ——— GG 3'

dATP + ddATP
 dCTP
 dGTP
 dTTP

GGCCA
 GGCCATCGTTGA

dATP
 dCTP + ddCTP
 dGTP
 dTTP

GGC
 GGCC
 GGCCATC

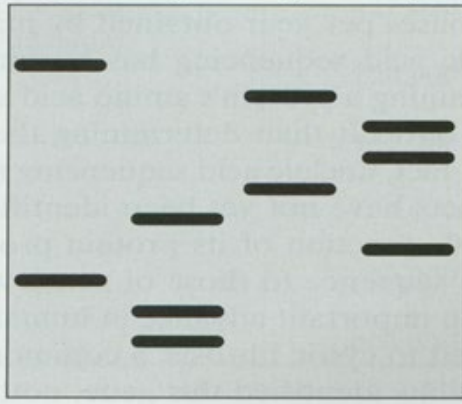
dATP
 dCTP
 dGTP + ddGTP
 dTTP

GGCCATCG
 GGCCATCGTTG

dATP
 dCTP
 dGTP
 dTTP + ddTTP

GGCCAT
 GGCCATCGT
 GGCCATCGTT

A C G T



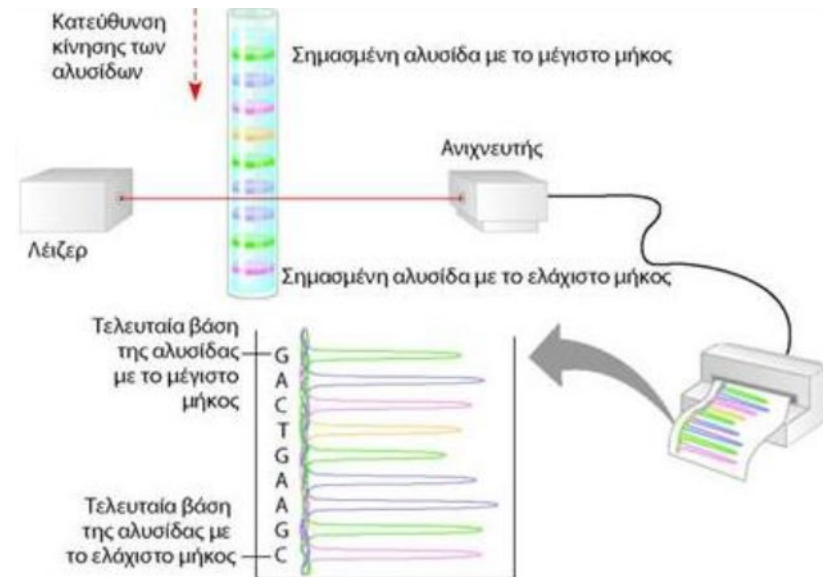
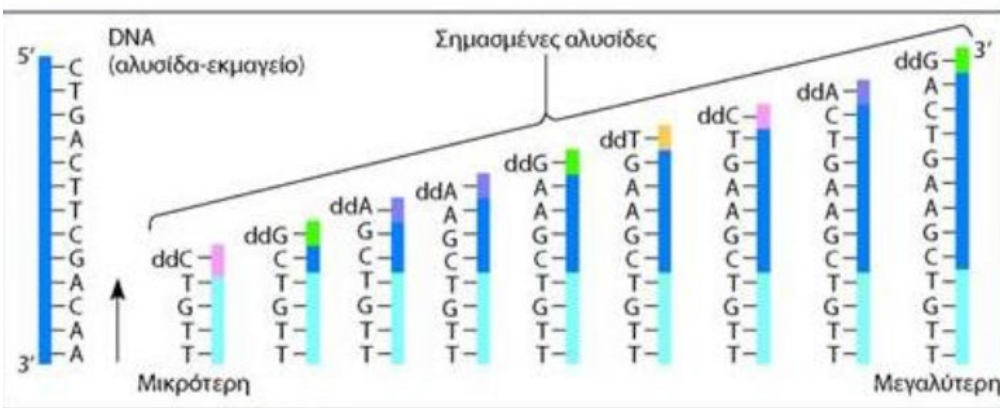
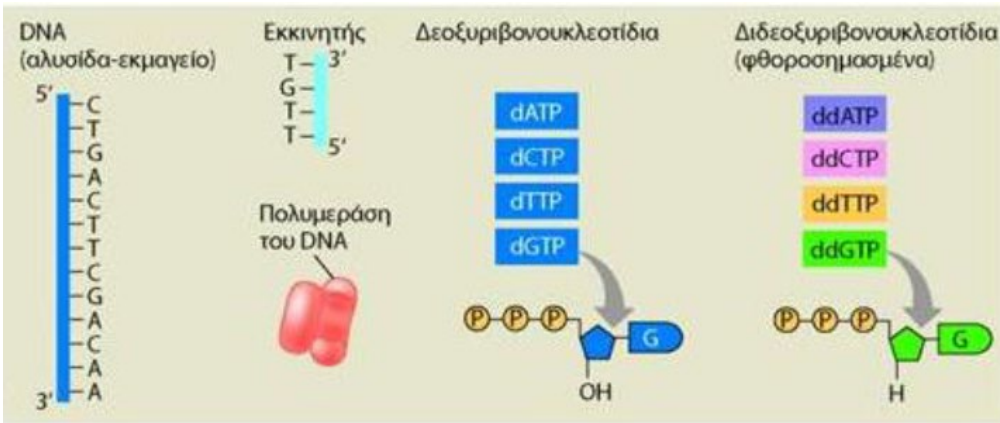
A 3'
 G
 T
 T
 G
 C
 T
 A
 C
 C 5'

Sequence complementary to template DNA

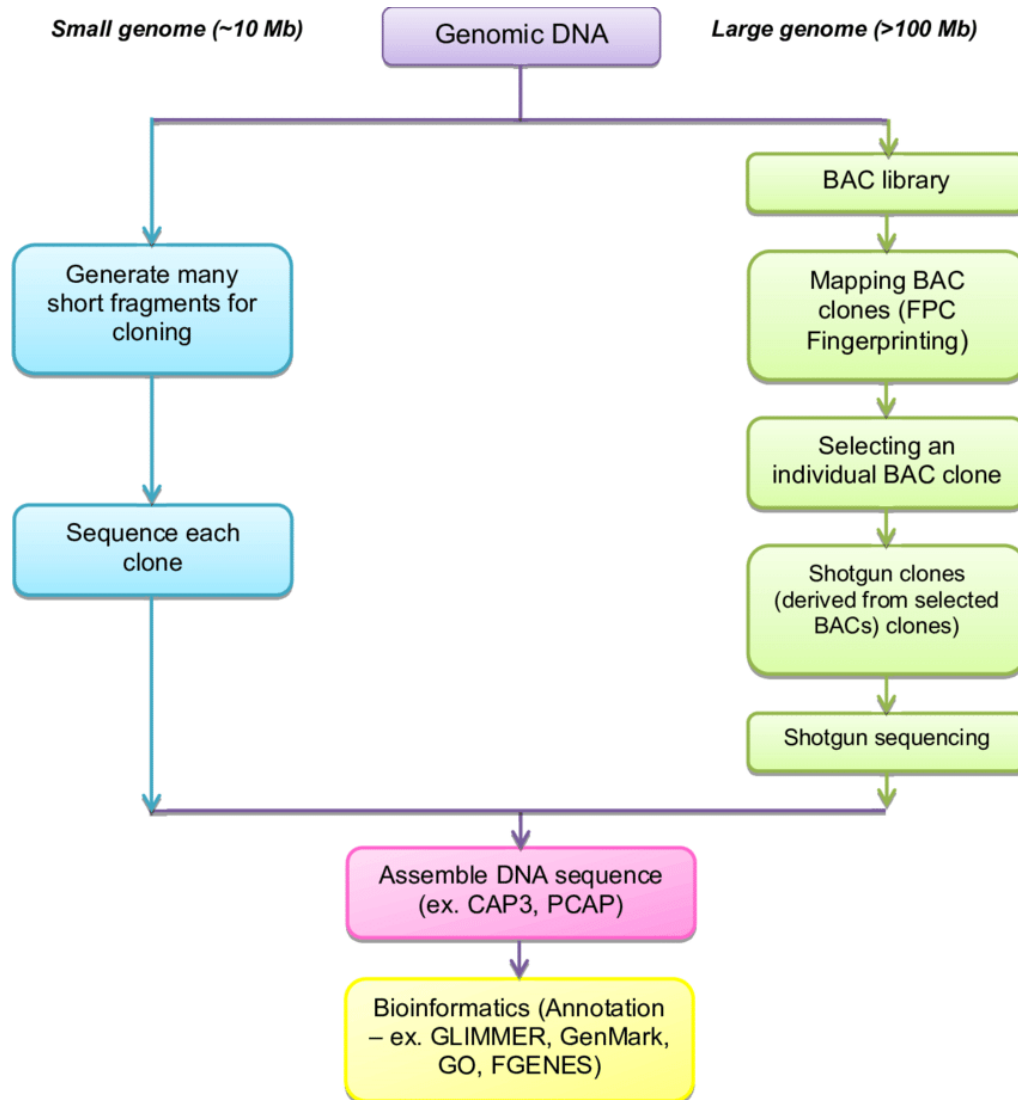
DYE TERMINATOR SEQUENCING

Το 1998 κυκλοφόρησε το πρώτο εμπορικά διαθέσιμο αυτόματο σύστημα ανάλυσης αλληλουχίας βάσεων.

- *Every ddNTP labelled with different fluorescent dye – all together in one reaction*
- *Separation by size in capillary – fluorescence detection*



Overview of whole-genome sequencing by conventional sequencing method



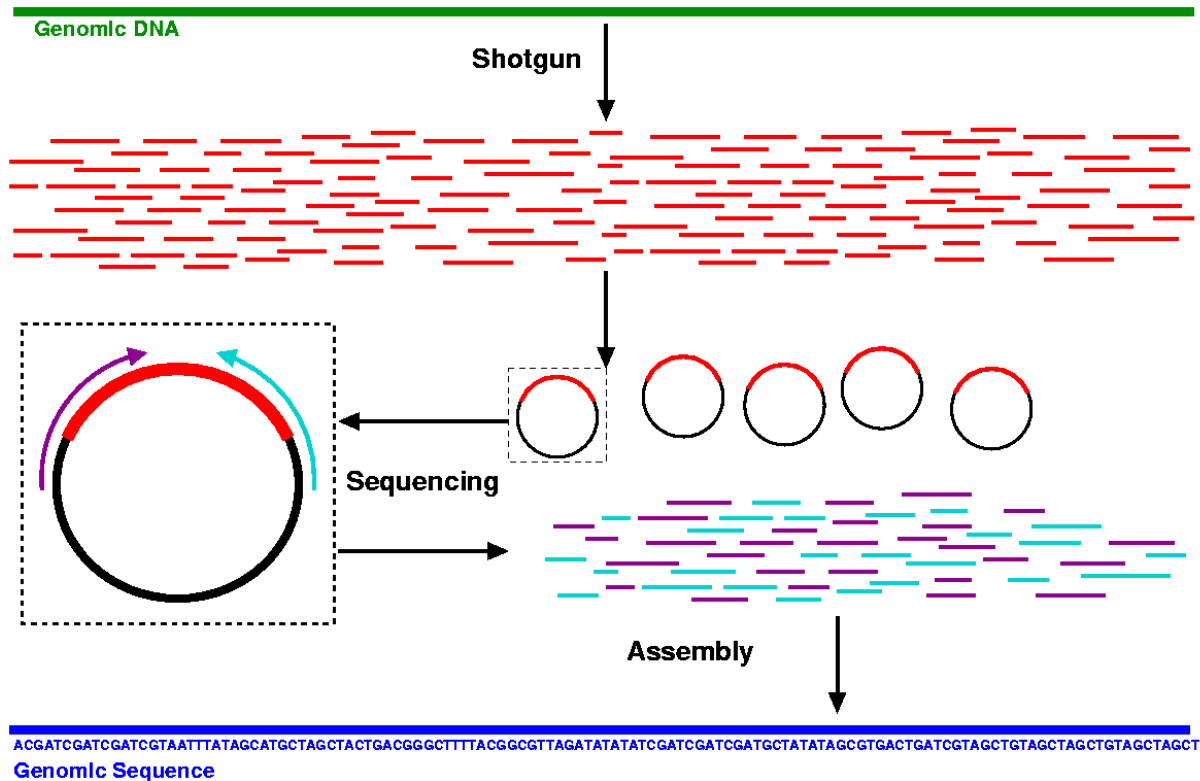
Η ανάλυση της αλληλουχίας ενός μεγάλου γονιδιώματος πλούσιου σε επαναλήψεις όπως το ανθρώπινο, μπορεί να γίνει με δύο τρόπους (International Human Genome Sequencing Consortium, 2001)

Human Genome Project

- Launched in 1989 – expected to take 15 years –
 - Competing Celera project launched in 1998
- Genome estimated to be 92% complete
 - 1st Draft released in 2000
 - “Complete” genome released in 2003
 - Sequence of last chromosome published in 2006
- Cost: ~\$3 billion
 - Celera ~\$300 million



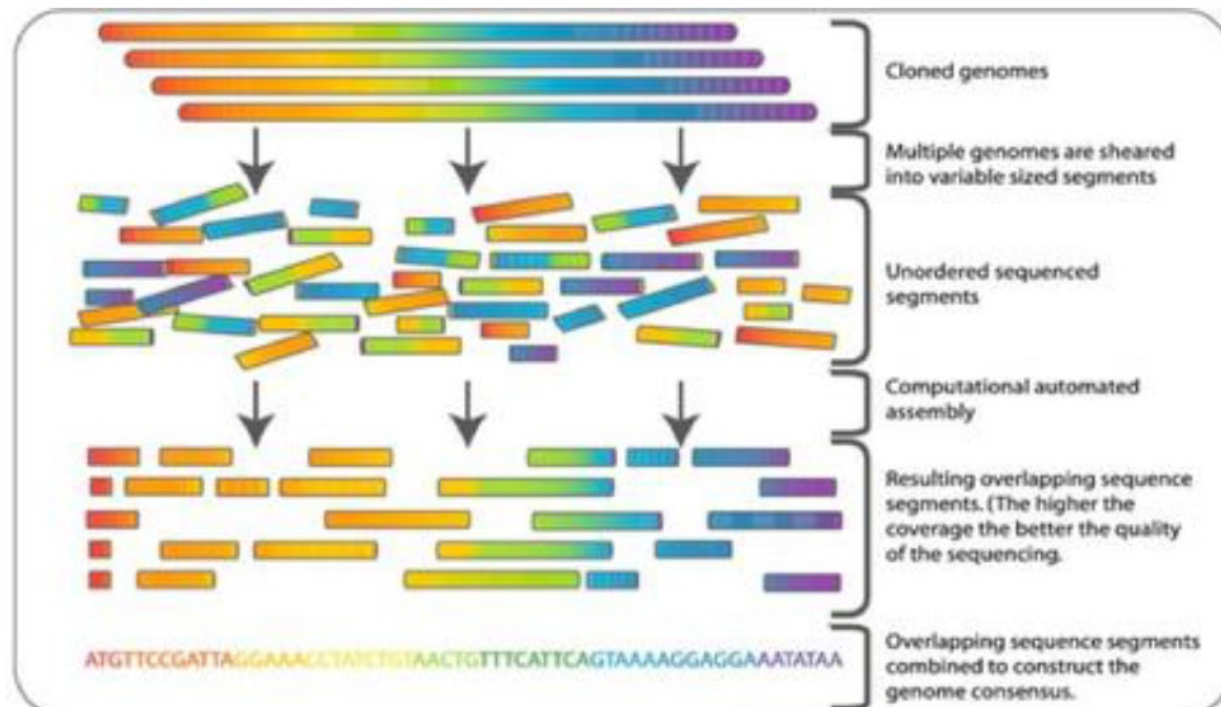
Shotgun: τα επικαλυπτόμενα τμήματα DNA που κλωνοποιούνται, προκύπτουν τυχαία από σπάσιμο ενθεμάτων BAC ή ολόκληρου του γονιδιώματος με τη βοήθεια υπερήχων ή μερικών πέψων με ένζυμα περιορισμού (όπως ακριβώς μια καραμπίνα χτυπάει με τα σκάγια της σε πολλά σημεία).



Ιεραρχική τμηματική» ανάλυση ("hierarchical shotgun sequencing" approach

- Αναφέρεται και ως «βασισμένη σε χάρτες» ("map-based"), «βασισμένη σε BAC» ("BAC-based") ή «κλώνο — κλώνο» ("clone - by - clone").
- Συνίσταται στη δημιουργία και οργάνωση ενός συνόλου κλώνων με μεγάλα ένθετα τμήματα (συνήθως 100-200 kb το καθένα) τοποθετημένων σε φυσικούς χάρτες που καλύπτουν το γονιδίωμα και κατόπιν στην τμηματική ανάλυση κατάλληλα επιλεγμένων κλώνων.
- Εξαιρετικά ακριβής, αλλά χρειάζεται χρόνο και είναι δαπανηρή

- Δημιουργείται μια γονιδιωματική βιβλιοθήκη σε BACs (συνήθως 100-200 kb το καθένα).
- Δημιουργείται χάρτης από επικαλυπτόμενα BACs.
- Επιλέγονται όσο το δυνατόν λιγότερα BACs για την πλήρη κάλυψη του γονιδιώματος.
- Τα BACs σπάζουν σε κομμάτια μεγέθους ~ 2 Kb που κλωνοποιούνται σε πλασμίδια.
- Ακολουθεί αλληλούχηση στα πλασμίδια μέχρις ότου επιτευχθεί μια 10πλάσια κάλυψη του BAC μέσω αλληλούχησης.



Μέθοδος τυχαίας προσπέλασης ολόκληρου γονιδιώματος», (whole - genome shotgun) από τον J. Craig Venter.

Χρησιμοποιήθηκε από την ιδιωτική εταιρεία Celera: Ολόκληρο το γονιδίωμα σπάζει τρεις φορές για να κατασκευαστεί:

- Μια πλασμιδιακή βιβλιοθήκη με ενθέματα ~ 2 Kb. Χρησιμοποιείται για 6πλάσια κάλυψη του γονιδιώματος
- Μια πλασμιδιακή βιβλιοθήκη με ενθέματα ~ 10 Kb. Χρησιμοποιείται για 3πλάσια κάλυψη του γονιδιώματος
- Μια BAC βιβλιοθήκη με ενθέματα ~ 200 Kb. Χρησιμοποιείται για μονή κάλυψη του γονιδιώματος

Χρησιμοποιούνται ενθέματα διαφορετικών μεγεθών επειδή βοηθάνε στην συναρμολόγηση τμημάτων μεταξύ των οποίων υπάρχουν κενά.

Ολική Γονιδιωματική Στρατηγική



Λογική συναρμολόγησης γονιδιωματικών διαβασμάτων από βιβλιοθήκες διαφορετικών ενθεμάτων: Στην κάτω σειρά φαίνονται οι συναρμολογήσεις από τα ενθέματα των μικρών βιβλιοθηκών (2 και 10 kb) οπότε έχουν προκύψει δύο ανεξάρτητα contigs (A και B). Παρ' όλο που αυτά δεν έχουν ενωθεί σε ένα contig η αλληλεπικάλυψη του καθενός με τα δύο άκρα από τον κλώνο 200 kb (πάνω σειρά) μας επιτρέπει να καταλάβουμε ότι αποτελούν μέρος ενός συνεχόμενου χρωμοσωματικού τμήματος, δηλαδή μας επιτρέπουν την καλύτερη συναρμολόγηση.

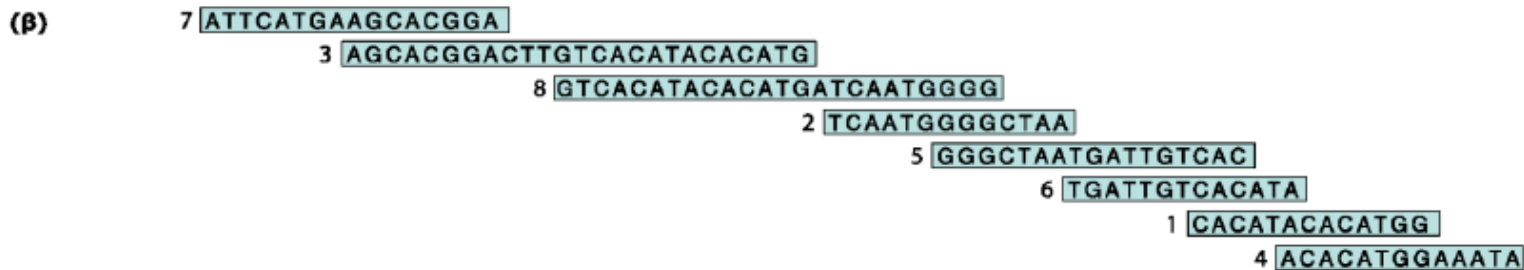


Η ολοκληρωμένη αλληλουχία ενός τμήματος DNA προσδιορίζεται μέσω της συναρμολόγησης επικαλυπτόμενων αναγνώσεων αλληλουχίας.

(α) Αναγνώσεις αλληλουχίας

- Ανάγνωση 1 **CACATACACATGG**
- Ανάγνωση 2 **TCAATGGGGCTAA**
- Ανάγνωση 3 **AGCACGGACTTGTACACATACACATG**
- Ανάγνωση 4 **ACACATGGAAATA**
- Ανάγνωση 5 **GGGCTAATGATTGTCAC**
- Ανάγνωση 6 **TGATTGTCACATA**
- Ανάγνωση 7 **ATTCATGAAGCACGGA**
- Ανάγνωση 8 **GTCACATACACATGATCAATGGGG**

↓
Συναρμολόγηση αναγνώσεων αλληλουχίας με τη βοήθεια υπολογιστή

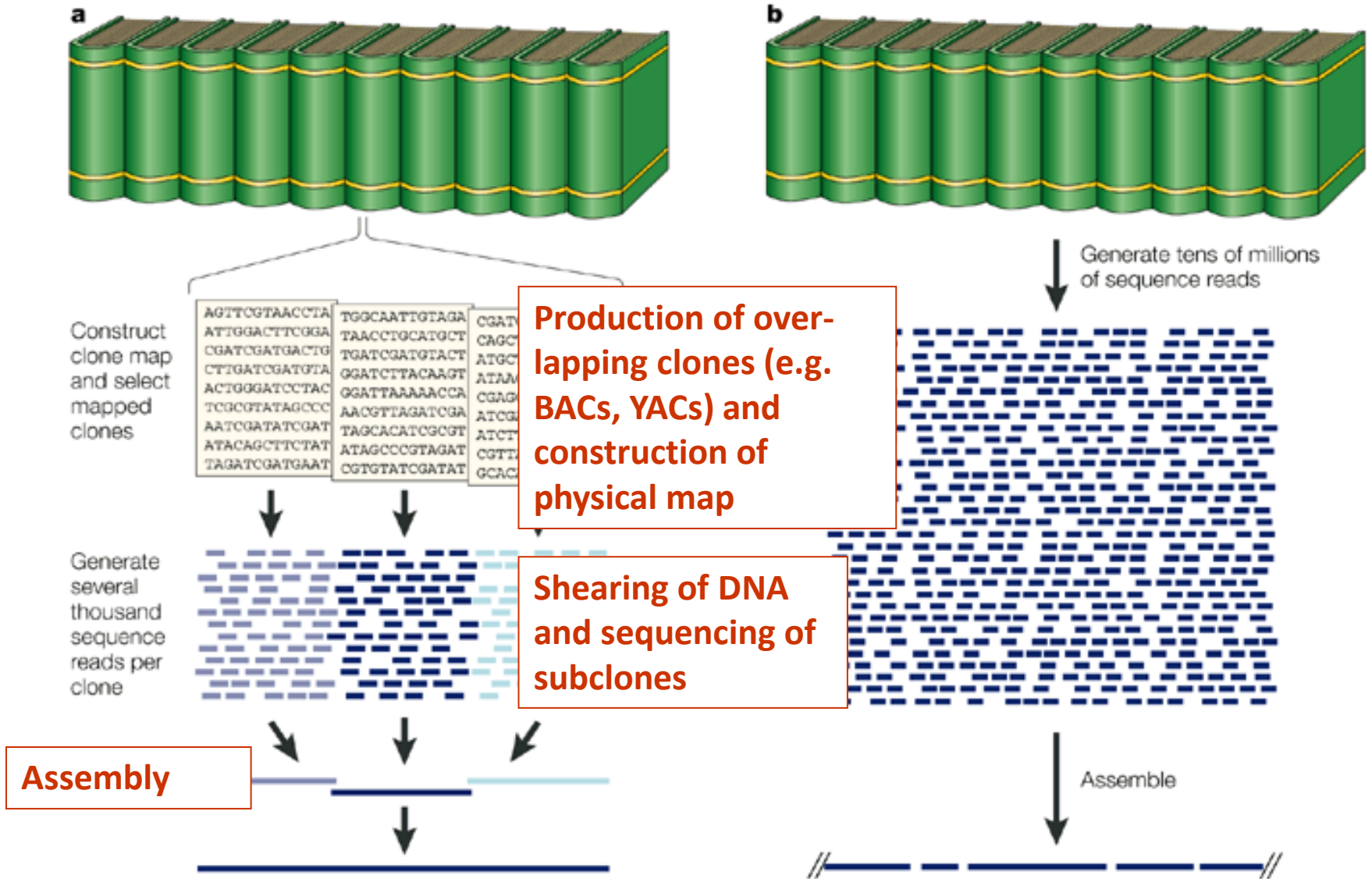


↓
Συναρμολογημένη αλληλουχία



Hierarchical shotgun sequencing

Whole-genome shotgun sequencing



Σύγκριση Στρατηγικών

- ❑ Μειονεκτήματα και των 2 μεθόδων είναι ότι πάντα υπάρχουν ετεροχρωματινικές συμπυκνωμένες περιοχές που είναι αδύνατο να κλωνοποιηθούν, και κάποιες κλωνοποιημένες περιοχές ανασυνδυάζονται ή χάνουν τμήματα τους
- ❑ Με την πάροδο του χρόνου, αποδείχθηκε ότι τα περισσότερα ερευνητικά κέντρα καταφεύγουν πια στην WGS προσέγγιση που είναι πιο γρήγορη και πιο φθηνή.

Οι ανάγκες και τα ερωτήματα αυξάνονται:

Χρειάζεται να γίνει γνωστή η λειτουργία 2500 γονιδίων για να κατανοηθεί το προκαρυωτικό γονιδίωμα

Η ανάγκη για συνολικές αναλύσεις οδήγησε στην ανάπτυξη ισχυρών **πλατφόρμων (high-throughput platforms)**.

Μια γονιδιωματική πλατφόρμα υψηλής ανάλυσης περιλαμβάνει:

- αυτόματα μηχανήματα για παραγωγή τμημάτων DNA για αλληλούχηση,
- αυτόματο αναλυτή αλληλουχίας
- και πλήθος υπολογιστικών μεθόδων για να συλλεχθούν και να επεξεργασθούν τα δεδομένα

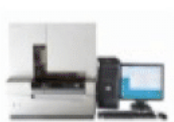
Αλληλούχηση επόμενης γενιάς (Next generation sequencing-NGS)

- Τα τελευταία χρόνια έχουν πραγματοποιηθεί επαναστατικές εξελίξεις στις τεχνολογίες της αλληλούχησης του DNA με την χρήση των τεχνικών αλληλούχησης επόμενης γενιάς. Οι μέθοδοι NGS επιτρέπουν την αλληλούχηση εκατομμυρίων βάσεων σε ένα γύρο, σε ένα κλάσμα του κόστους σε σχέση με τις παραδοσιακές αλληλουχήσεις με τη μέθοδο Sanger.
- Σήμερα, είναι εφικτό να αλληλουχηθούν τα περισσότερα φυτικά γονιδιώματα (με εξαίρεση εκείνα με πολύ μεγάλο και πολύπλοκο γονιδίωμα) με ένα σχετικά μικρό προϋπολογισμό, συνδυάζοντας τις τεχνολογίες Sanger και τις NGS.

First generation

Second generation (next generation sequencing)

Third generation



Sanger sequencing
Maxam and Gilbert
Sanger chain termination

454, Solexa,
Ion Torrent,
Illumina

PacBio
Oxford Nanopore

Infer nucleotide identity using dNTPs,
then visualize with electrophoresis

High throughput from the
parallelization of sequencing reactions

Sequence native DNA in real time
with single-molecule resolution

500–1,000 bp fragments

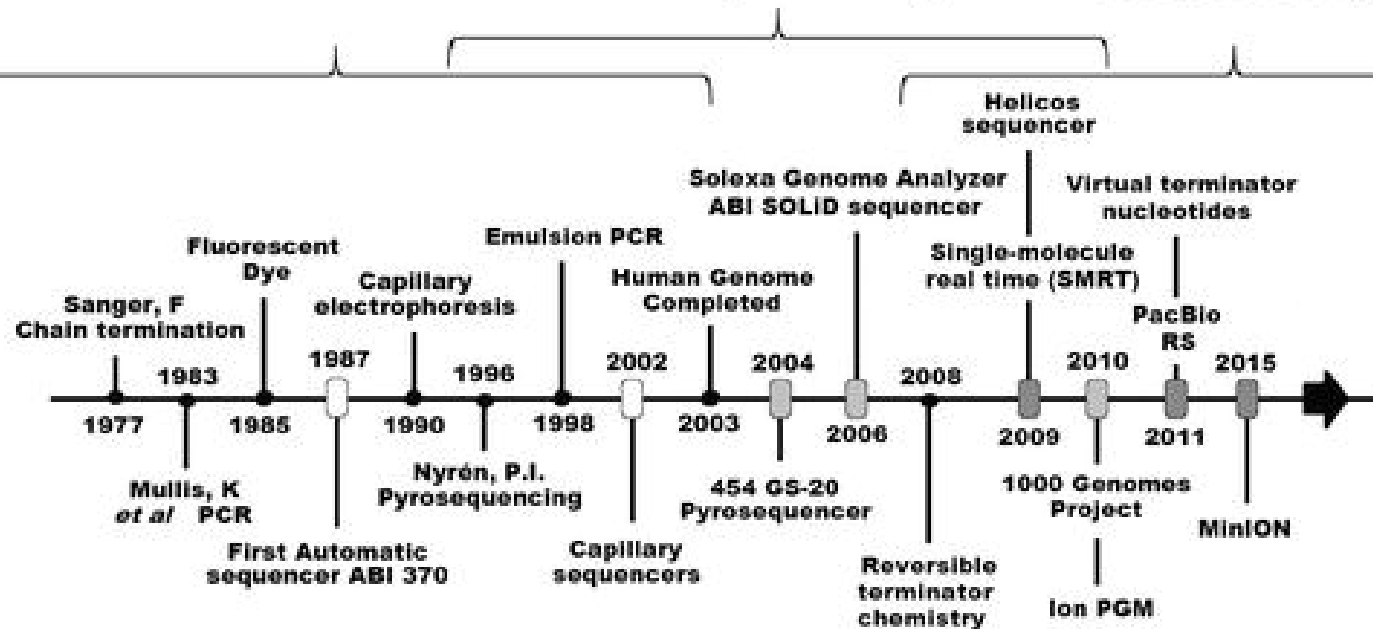
~50–500 bp fragments

Tens of kb fragments, on average

Short-read sequencing

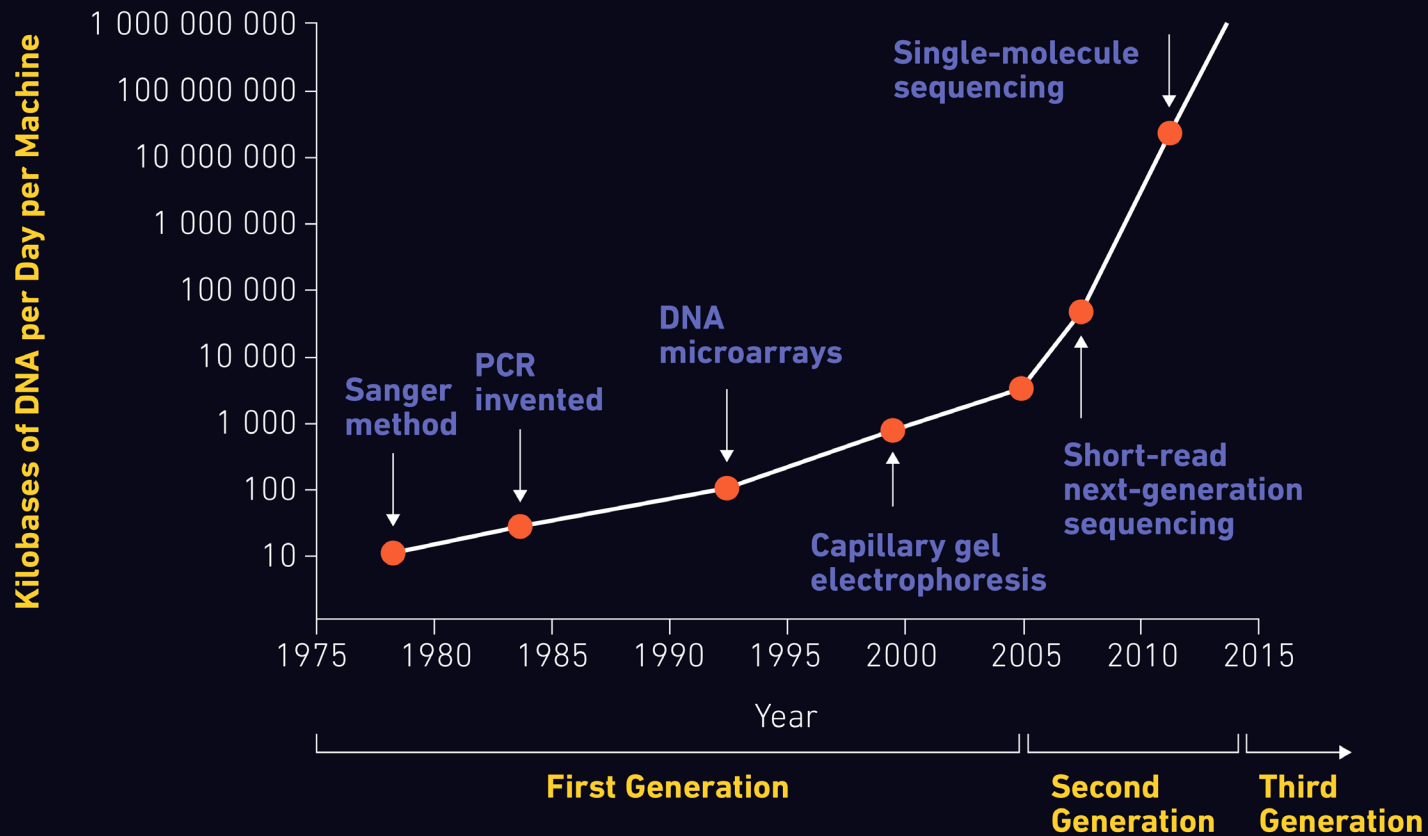
Long-read sequencing

1st Generation DNA sequencing **2nd Generation Massive Parallel sequencing** **3rd Generation Single Molecule sequencing**

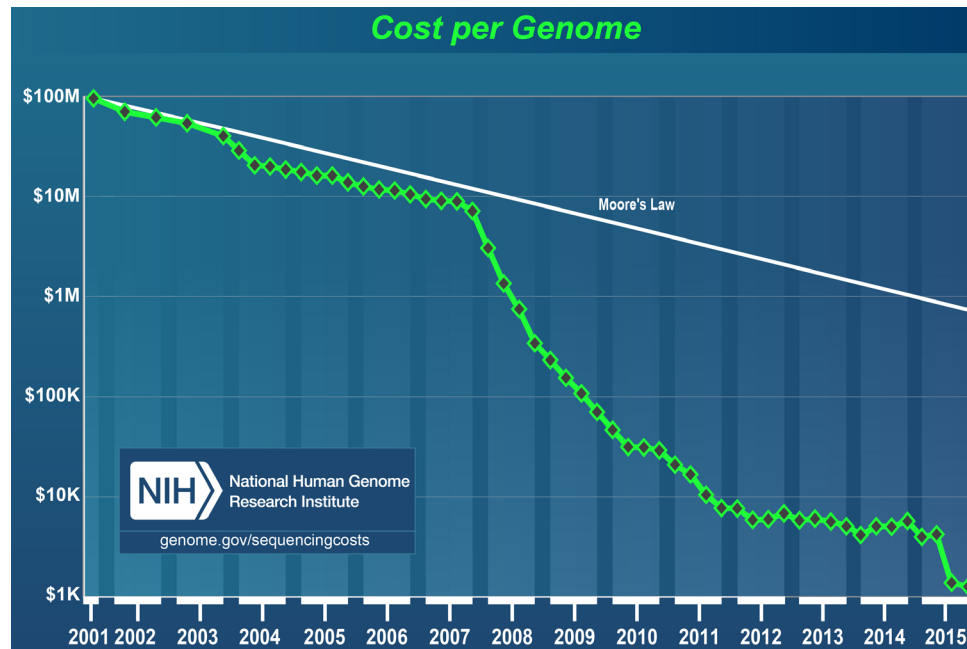
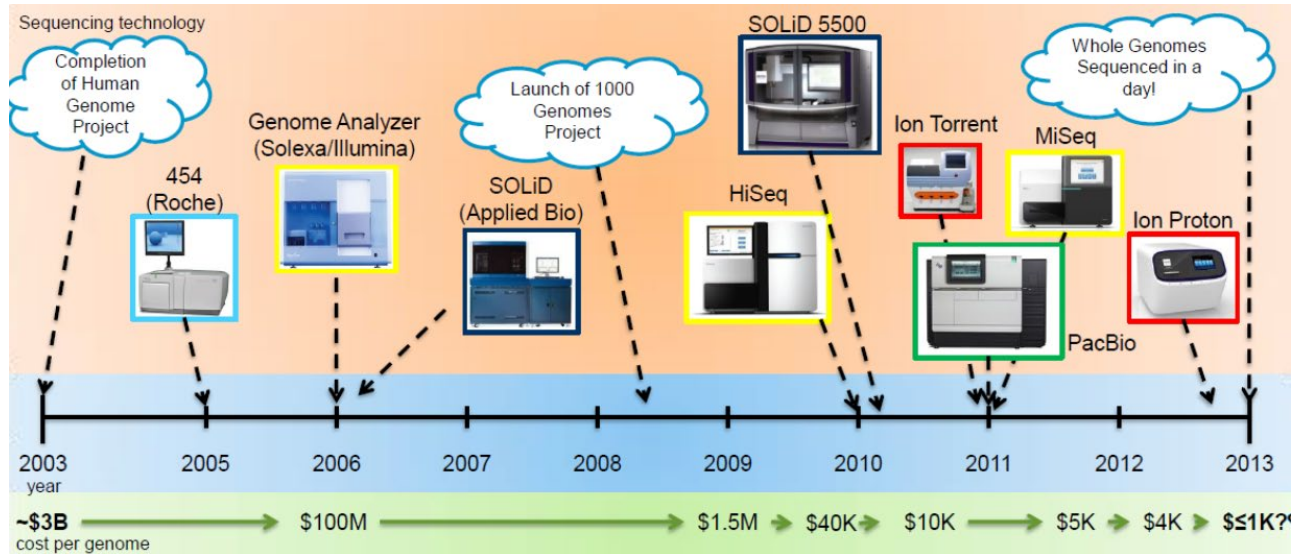


Cost per Raw Megabase of DNA Sequence in US Dollar

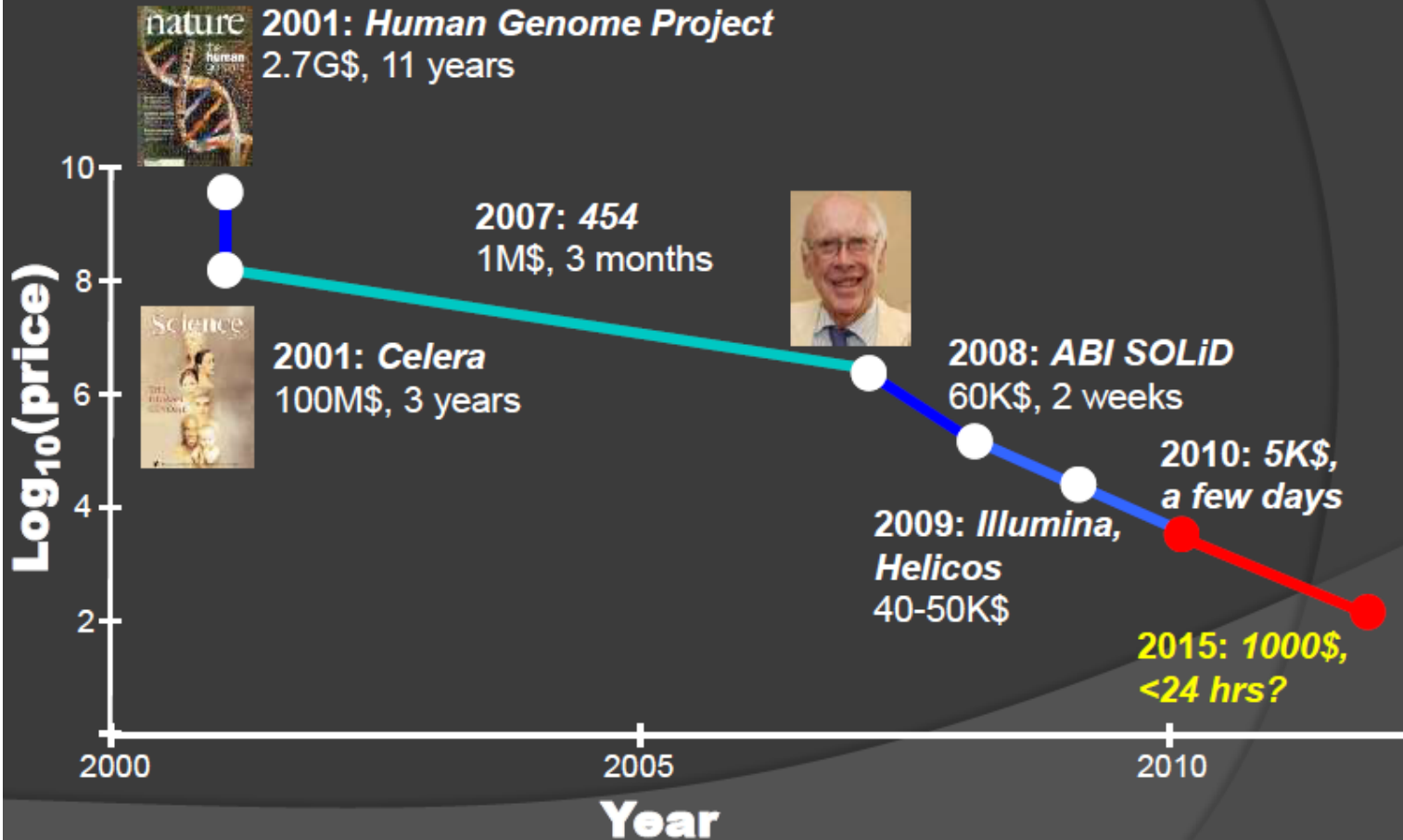




Η εξέλιξη της αυτοματοποιημένης παράλληλης αλληλούχησης



Sequencing the Human Genome



First generation sequencing and Next (second) Generation Sequencing

Ένα τμήμα DNA (Μέθοδος Sanger)

➔ Πολλαπλά τμήματα DNA (NGS)

The Nobel Prize in Chemistry 1980



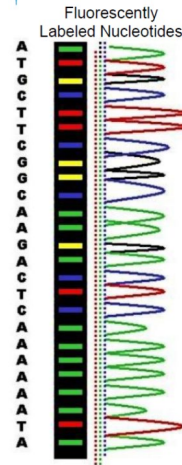
Paul Berg
Prize share: 1/2



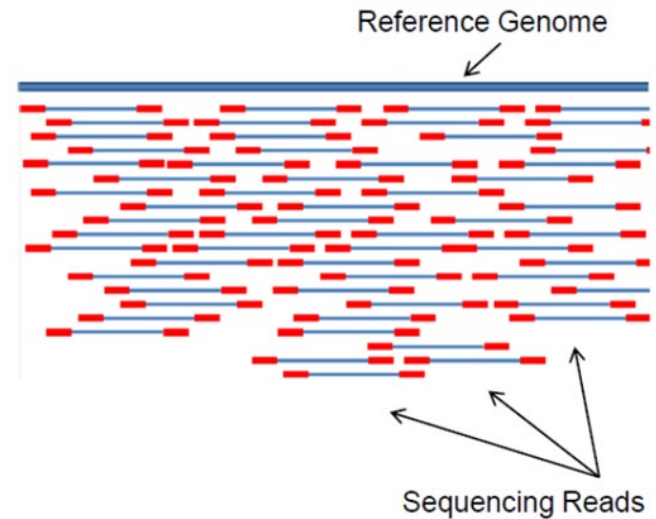
Walter Gilbert
Prize share: 1/4



Frederick Sanger
Prize share: 1/4



Fluorescently labeled nucleotides of many different DNA fragments being sequenced in parallel



Cell

All Content

Cell All cell.com

Explore Online Now Current Issue Archive Journal Information For Authors

< Previous Article Volume 155, Issue 1, p27–38, 26 September 2013

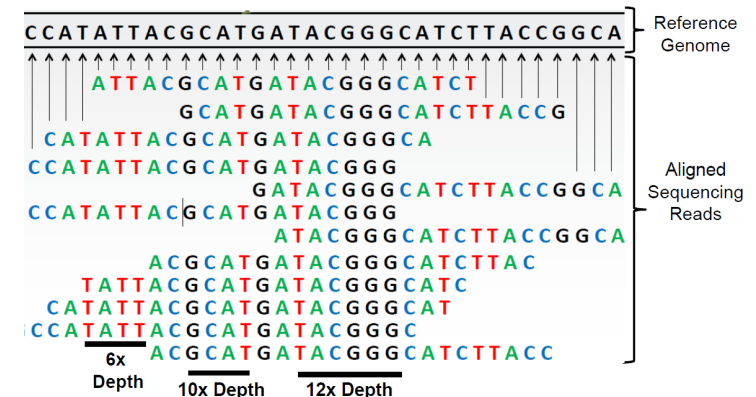
REVIEW

The Next-Generation Sequencing Revolution and Its Impact on Genomics

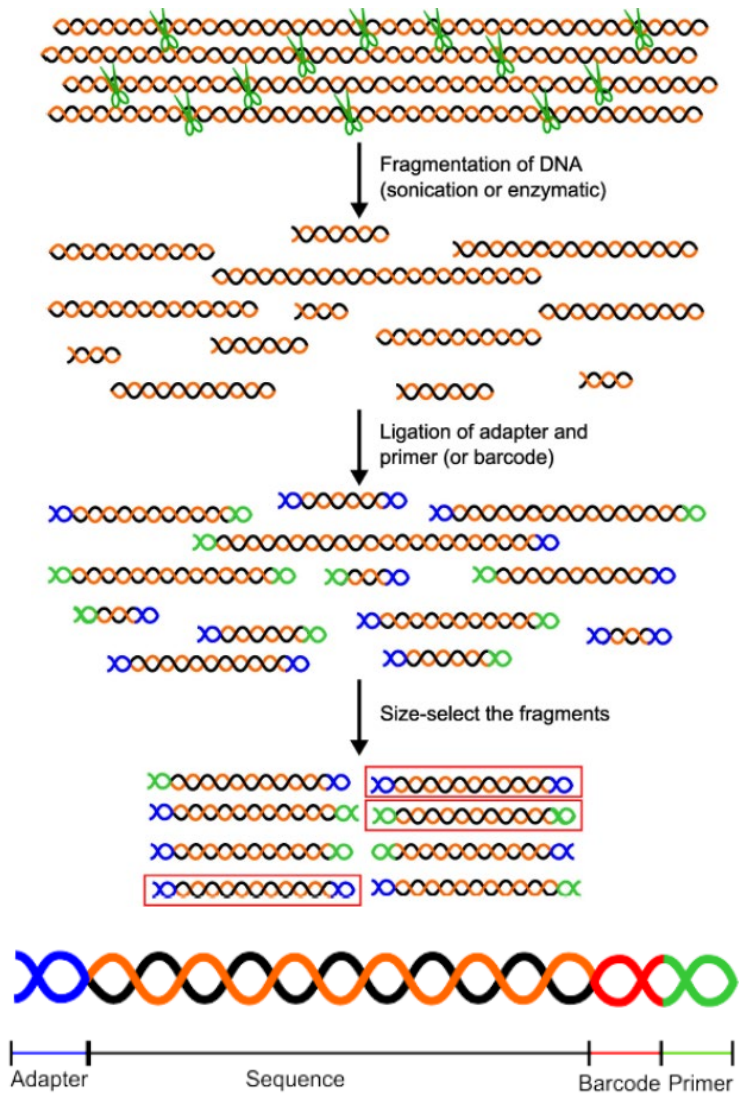
Daniel C. Koboldt, Karyn Meltz Steinberg, David E. Larson, Richard K. Wilson, Elaine R. Mardis

Open Archive PlumX Metrics

DOI: <https://doi.org/10.1016/j.cell.2013.09.006> | CrossMark



Χαρακτηριστικά της NGS



- DNA ή RNA
- **Κατακερματισμός σε μικρότερα τμήματα**
- Καθήλωση σε ειδικά υλικά (σφαιρίδια και chips με ειδικές υποδοχές)
- **Barcoding** προκειμένου να γνωρίζουμε την ταυτότητα κάθε δείγματος, ή πληθυσμού μορίων
- Ενίσχυση της αλληλουχίας (απαραίτητη στα RNAs τα οποία δεν βρίσκονται σε αφθονία και σε μελέτες έκφρασης)
- Αλληλούχηση, επεξεργασία και ανάκτηση δεδομένων online (over the cloud) ή μέσω ειδικών user-friendly λογισμικών
- **Ταχύτητα** (1987: ABI Automated Sequencer 4800 bp/day, Σήμερα: το γονιδίωμα ενός βακτηρίου σε 3 ώρες με 99.6% ακρίβεια)
- **Ακρίβεια (99.9%** με προϋποθέσεις σε μελέτες DNA αλληλούχησης ή επιπέδων γονιδιακής έκφρασης)
- **De novo δεδομένα** (και άρα νέοι δείκτες)
- **Πληθώρα δεδομένων!!!** (Πολλά Gb ακόμα και για μικρού εύρους αναλύσεις)

WHOLE GENOME SEQUENCING

1 Break genome into large fragments and clone

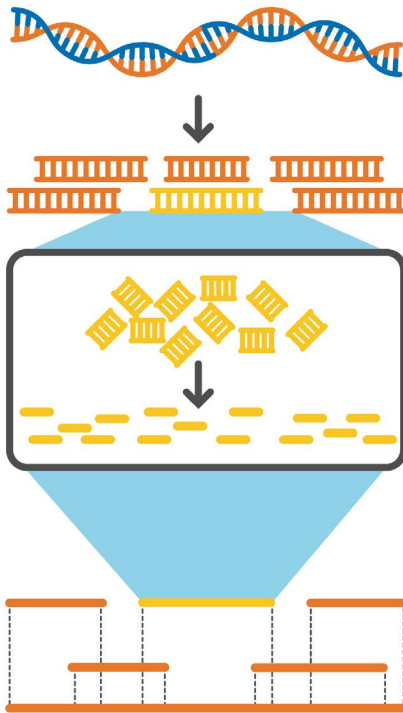
2 Break individual clone into small fragments

3 Generate thousands of sequence reads

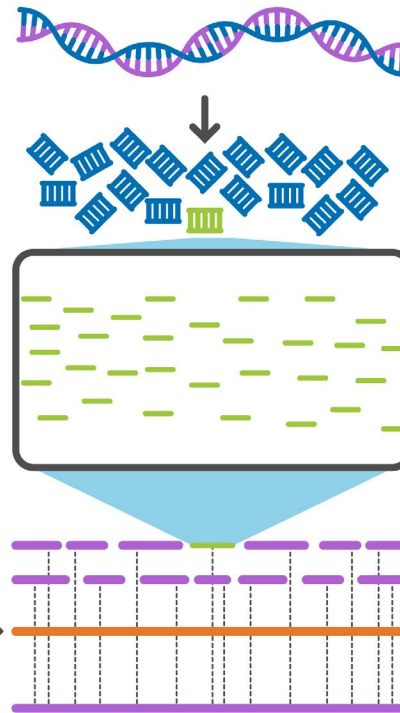
4 Assemble sequence reads for each clone

Reference genome

Reference Genome



Individual Genome



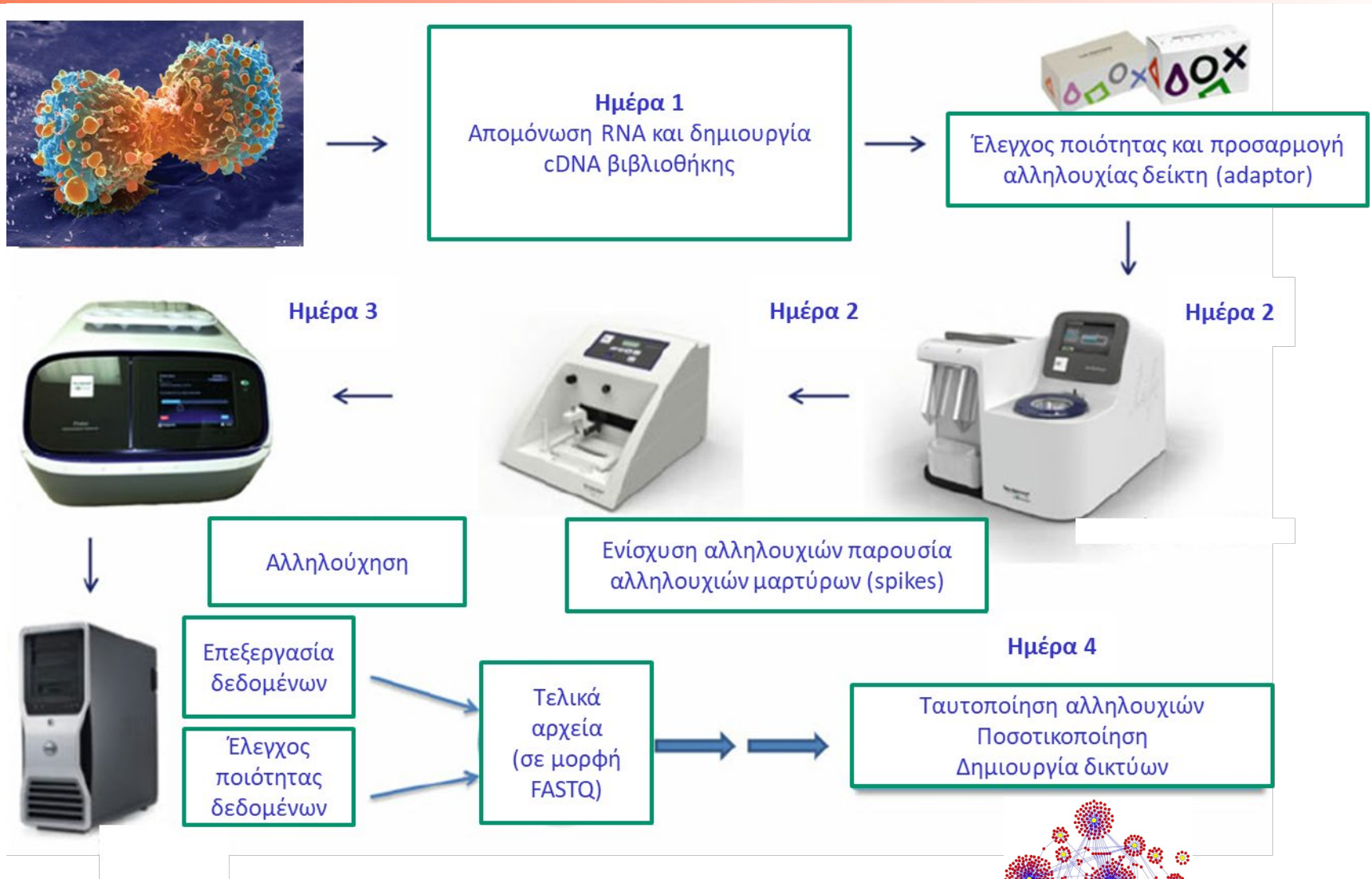
1 Break genome into small fragments

2 Generate millions of sequence reads

3 Align sequence reads into a reference genome

Individual genome

Διάγραμμα ροής ανάλυσης με NGS

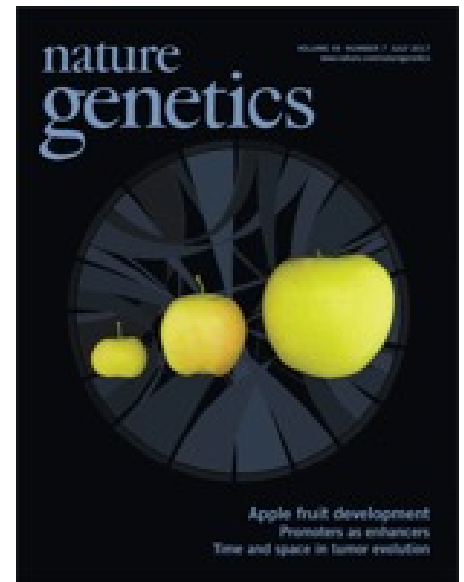


Σύγκριση των κύριων χαρακτηριστικών της συμβατικής αλληλούχησης (μέθοδος Sanger) και μερικών τεχνολογιών NGS

Τεχνολογία	Μήκος διαβάσματος (bp)	Mbp σε ένα γύρο	Κόστος (\$/Mbp)
Sanger	1000	0.001	3000.00
454 Roche	450	450	66.00
Illumina Hi-Seq2000	100	270000	0.07
Solid 5500xl	50	270000	0.07

Αλληλουχηθέντα φυτικά γονιδιώματα τα τελευταία χρόνια

Φυτικό είδος	Μέγεθος γονιδιώματος	Έτος πλήρους αλληλούχισης
<i>Beta vulgaris</i> (ζαχαρότευτλο)	714-758 Mbp	2013
<i>Prunus persica</i> (ροδάκινο)	265 Mbp	2013
<i>Musa acuminata</i> (μπανάνα)	523 Mbp	2012
<i>Citrus clementina</i> (κλημεντίνια)		2013
<i>Pyrus communis</i> (αχλάδι)		2013
<i>Nicotiana glauca</i> (καπνός)	~ 3.6 Gbp	2013
<i>Solanum lycopersicum</i> (τομάτα)	~900 Mbp	2012
<i>Solanum tuberosum</i> (πατάτα)	~844 Mbp	2011
<i>Phaseolus vulgaris</i> (φασόλι)		2013
<i>Capsicum annuum</i> (πιπεριά)	~3.48 Gbp	2014
<i>Triticum aestivum</i> (σιτάρι)		2012



Ten Thousand Plant Genome Project

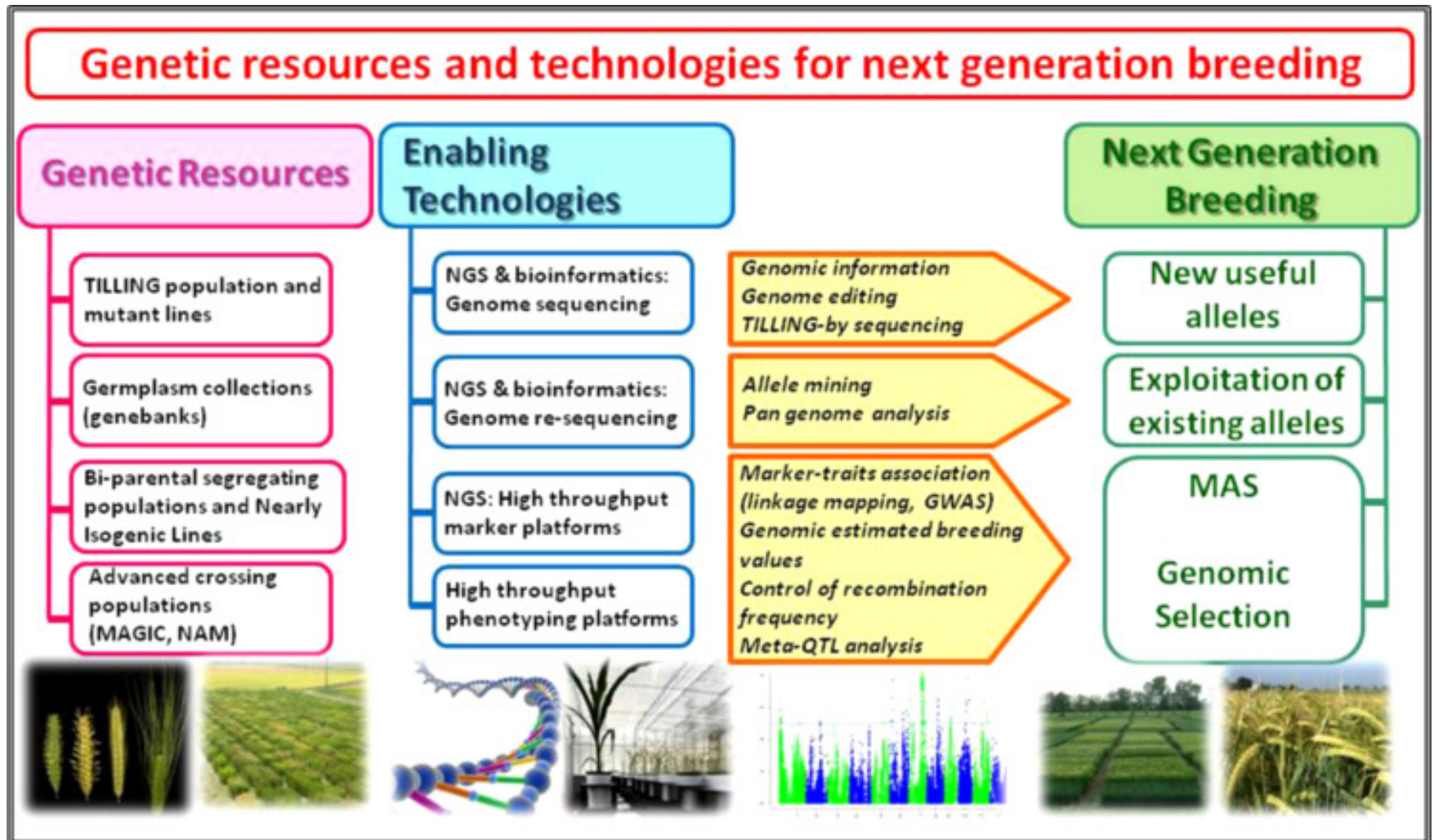
<https://en.genomics.cn/en-project-dzwyjs-6184.html>

The Ten Thousand Plant Genome Project (10KP) έχει ως στόχο την αλληλούχηση των γονιδιωμάτων 10,000 φυτών

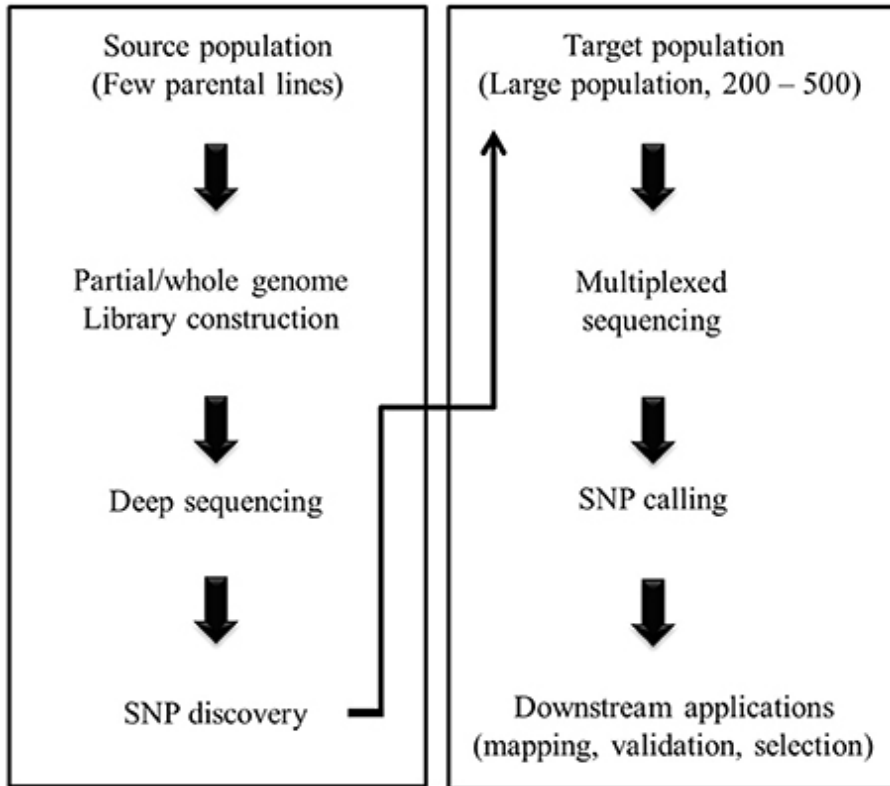
The Ten Thousand Plant Genome Project (10KP) was jointly initiated by BGI-Shenzhen (Institute of Life Sciences) and several authoritative experts in the field of botany during the 19th International Botany Conference in 2017.



Χρήση NGS στη Βελτίωση Φυτών



Χρήση NGS στη Βελτίωση Φυτών



GBS has been used in development of high density map of 20000 SNPs in wheat and 34000 SNPs in barley and to map QTLs for spike architecture and reduced plant height in barley

- Reference genomes: Περισσότερα από 100 γονιδιώματα καλλιεργουμένων φυτών
- Transcriptome research με χρήση RNA-sequencing (RNA-seq) – ανάλυση έκφραση γονιδίων, αποκάλυψη γονιδίων
- Επιγονιδιωματικές αναλύσεις: expression of small RNAs, and modification of chromatins (i.e. DNA methylation and adjustment of histone tails like in the cases of acetylation, methylation, ubiquitination, and phosphorylation).
- Εύρεση γενετικών δεικτών για MAS.
- GWAS για genomic selection.

Οι πλατφόρμες NGS που επικρατούν σε εθνικό και παγκόσμιο επίπεδο

DNA whole genome (ιοί, βακτήρια, μύκητες) Amplicon seq, ChiP seq, Methylation sites (epigenome)

RNA sequencing (whole transcriptome, exome sequencing, small and long non coding RNAs,

Χρήση σε πολλούς τομείς εκτός έρευνας και διάγνωσης (τρόφιμα, περιβαλλοντικά δείγματα

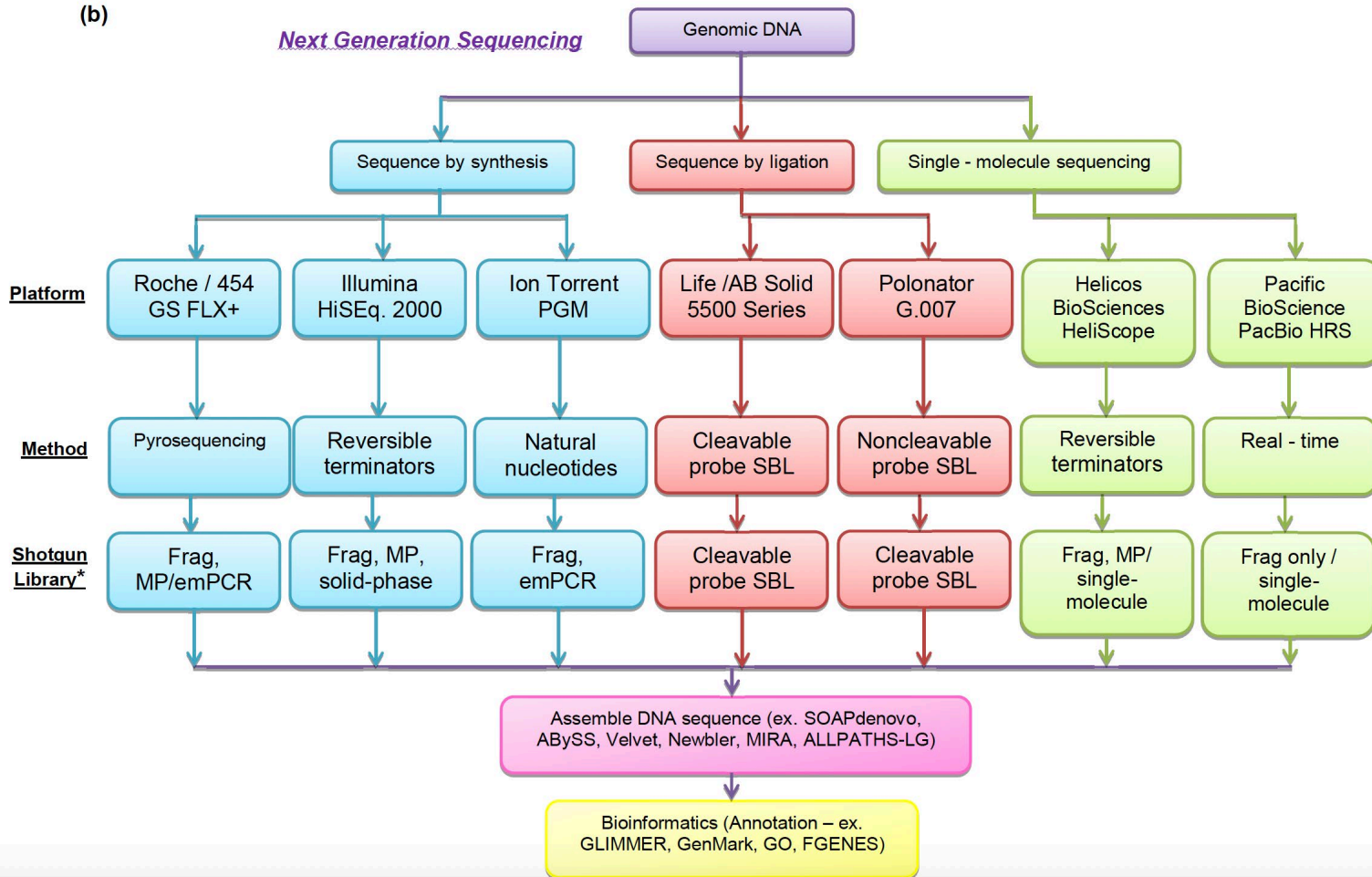
- Roche/454 FLX: 2004
- Illumina Solexa Genome Analyzer: 2006
- Applied Biosystems SOLiD™ System: 2007
- Applied Biosystems Ion Torrent: 2010
- Pacific Biosciences SMRT: 2010
- Helicos Heliscope™



Overview of whole-genome sequencing by next-generation sequencing method.

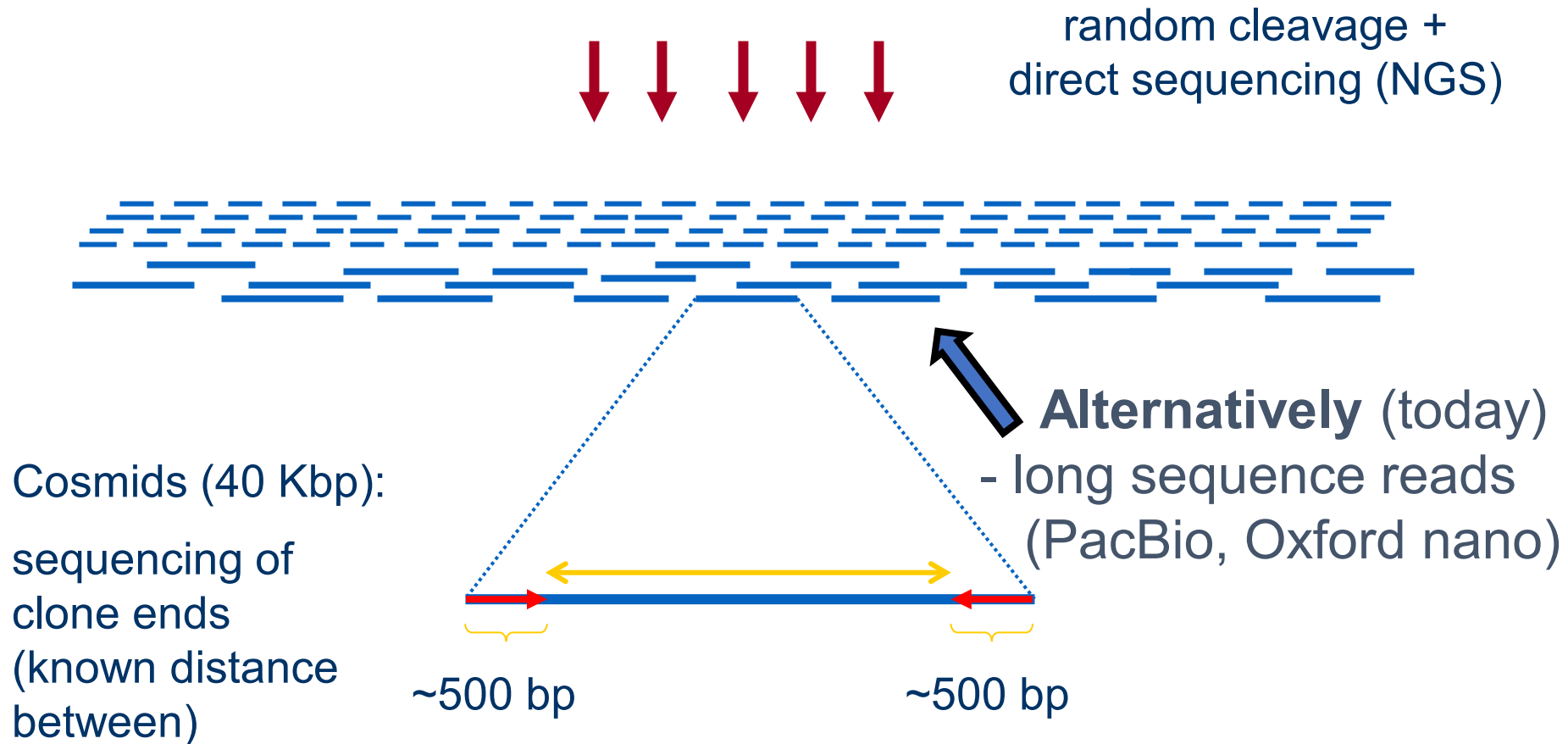
Van et al. (2013)

(b)



Shotgun sequencing

BAC/chromosome/whole genome

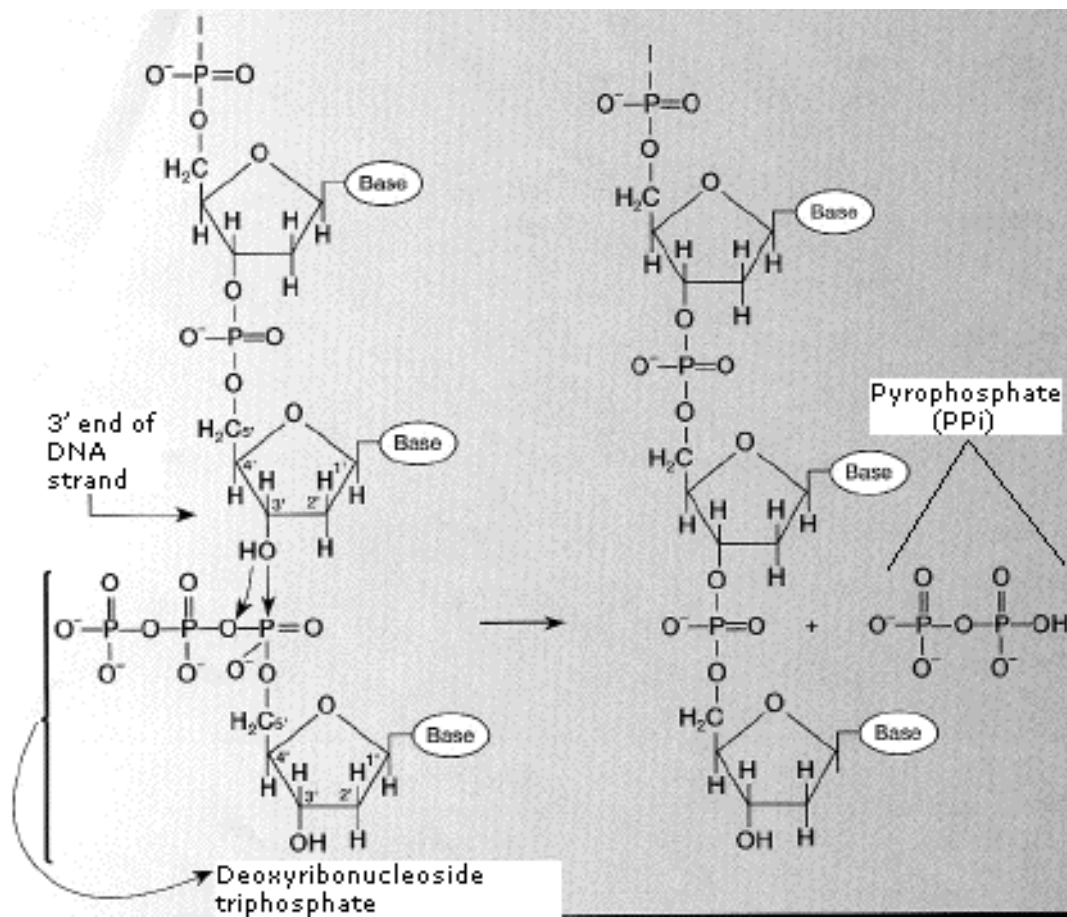


Pyrosequencing

Η ανάλυση της αλληλουχίας με Pyrosequencing είναι μια μέθοδος με υψηλή ακρίβεια η οποία βασίζεται στην αρχή “αλληλούχηση μέσω σύνθεσης” με την οποία παρακολουθούμε ποσοτικά την ενσωμάτωση των νουκλεοτιδίων από μια πολυμεράση σε πραγματικό χρόνο. Ο προσδιορισμός του Pyrosequencing βασίζεται στην ανίχνευση της παραγωγής φωτός όταν απελευθερώνεται ανόργανο πυροφωσφορικό.

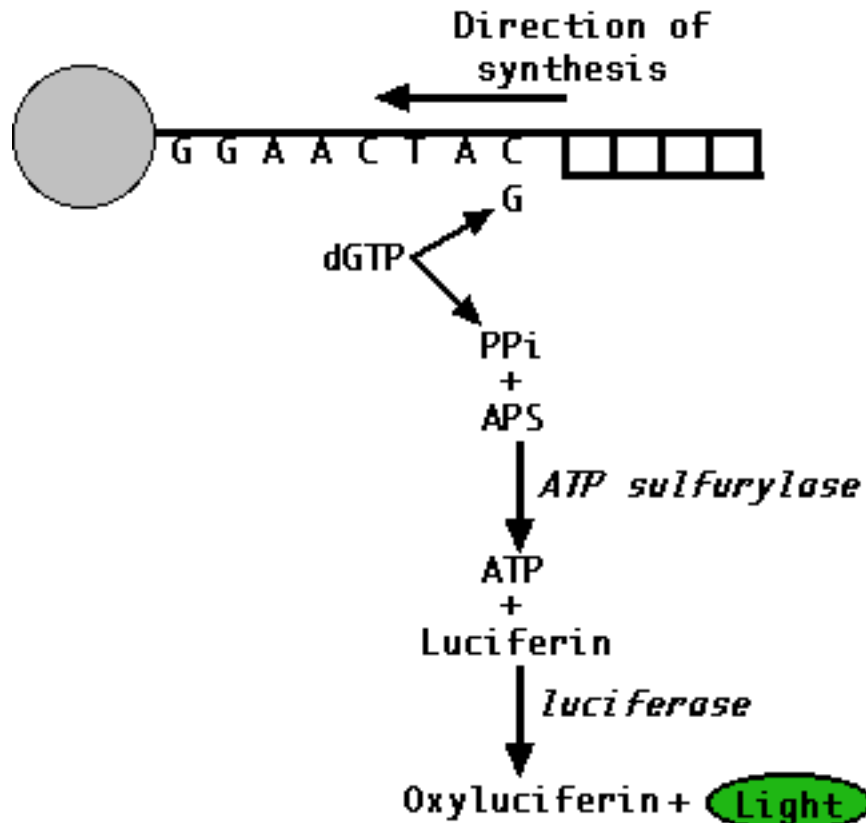
Κατά την σύνθεση του DNA όταν ένα dNTP προστίθεται στο 3' άκρο της εκτεινόμενης αλυσίδας του DNA τα δύο φωσφορικά στο άκρο του dNTP απελευθερώνονται εκλύοντας πυροφωσφορικό PPI.

Αν δεν υπάρχει προσθήκη dNTP δεν θα ελευθερωθεί και πυροφωσφορικό.

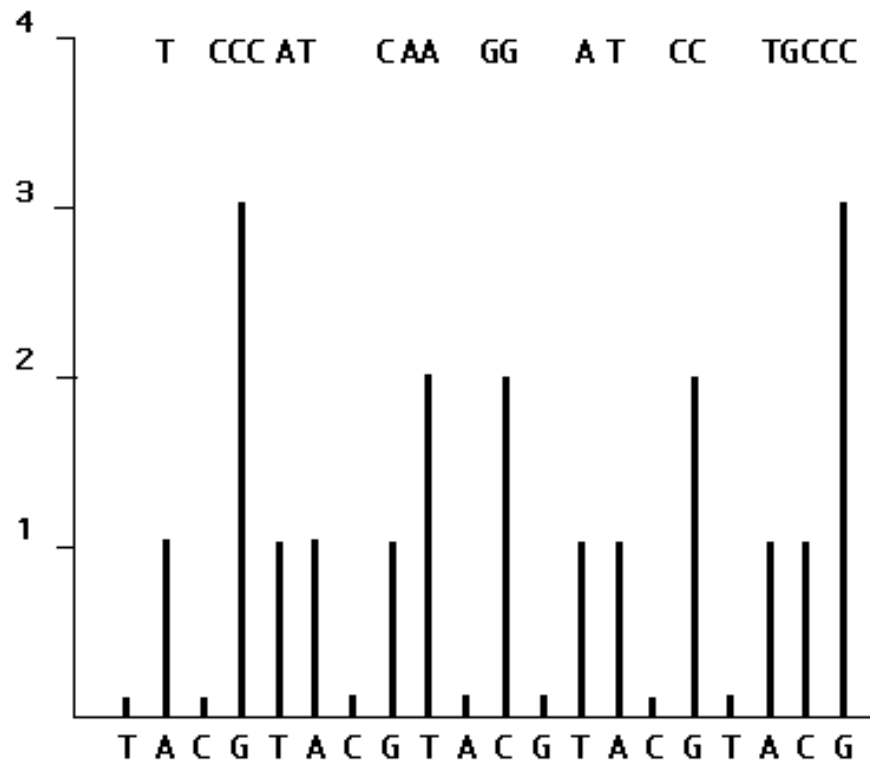


Με μια σειρά χημικών αντιδράσεων με το πυροφωσφορικό που έχει παραχθεί, εκλύεται ορατό φώς.

Το παραγόμενο φωτεινό σήμα είναι ανάλογο του αριθμού των dNTPs που προστίθενται. Η διαδικασία επαναλαμβάνεται με το επόμενο dNTP.



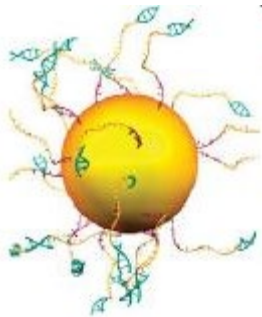
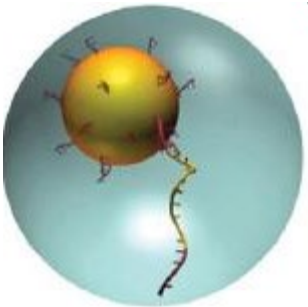
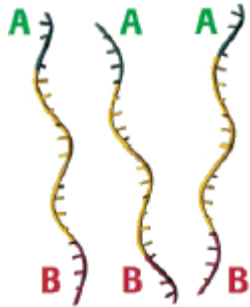
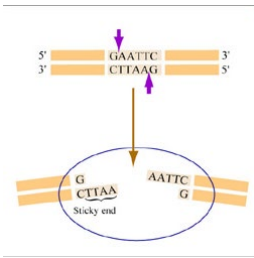
- Το παραγόμενο πυροδιάγραμμα από ένα κελί-φρεάτιο.



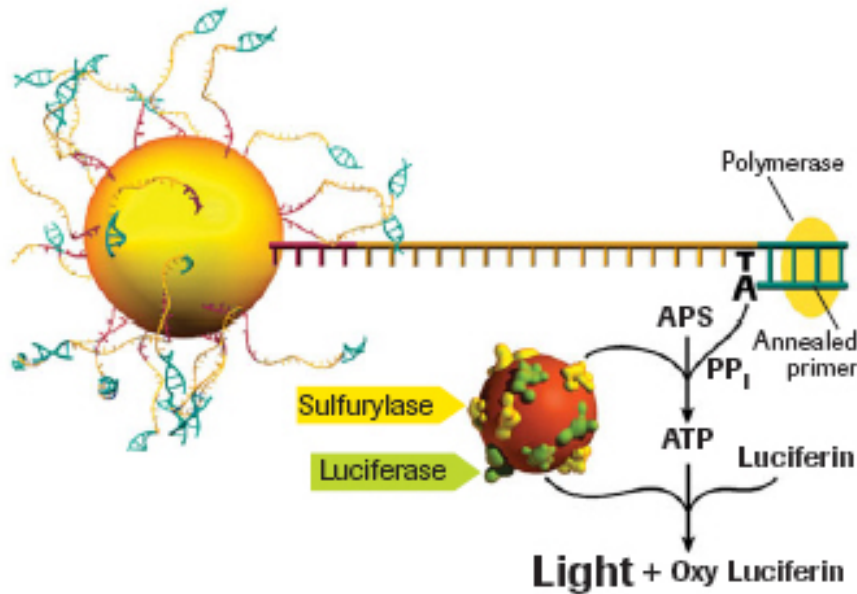
Roche / 454

- Τεμαχισμός DNA
- Σύνδεση μονόκλωνου DNA με αλληλουχίες αντάπτορες που επιτρέπουν τη σύνδεση σε μοναδική μπίλια 28 μm διαμέτρου
- Ενσωματώνονται σε σταγόνα λαδιού όπου υπάρχουν τα απαραίτητα υλικά για emPCR
- Οι σταγόνες λαδιού δεν ενώνονται - όχι επιμόλυνση
- Οι μπίλιες με το υπόστρωμα DNA τοποθετούνται μέσα σε σλάιντ φωτοευαίσθητων οπτικών ινών όπου υπάρχουν και τα απαραίτητα ένζυμα για αλληλούχηση
- Με κάθε προσθήκη ενός νουκλεοτιδίου υπάρχει φωτοευαίσθητος αισθητήρας που διαβάζει το νουκλεοτίδιο σε 1 από τους 1,6 εκατομ. αντιδραστήρες!

454 FLX, Roche



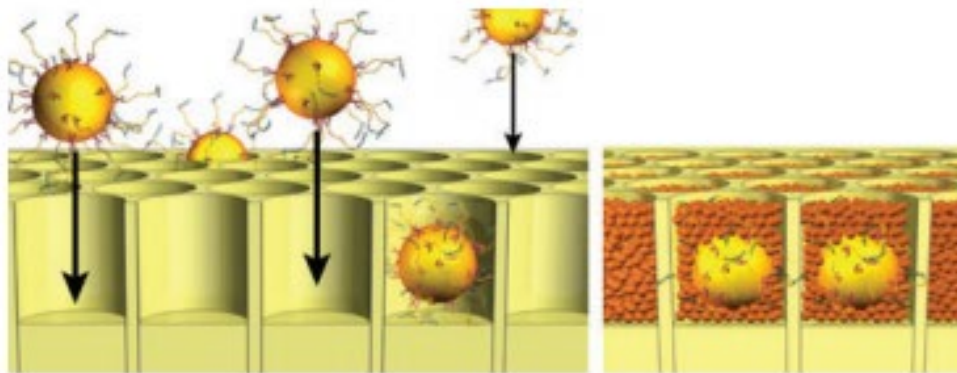
- Η μέθοδος της πυροαλληλούχισης εξελίχθηκε για να γίνει αποδοτικότερη και λιγότερο δαπανηρή.
- Το DNA κόβεται σε τμήματα των 300-800bp απομακρύνοντας τα εξέχοντα νουκλεοτίδια
- Στα άκρα συνδέονται υποδοχείς και μετατρέπονται σε μονόκλωνα.
- Το κάθε μονόκλωνο DNA προσκολλάται στην επιφάνεια ενός μικροσφαιριδίου 28μm. Μόνο ένα sDNA πρέπει να υπάρχει σε κάθε μικροσφαιρίδιο.
- Με PCR πολλαπλασιάζεται το DNA. Κάθε μικροσφαιρίδιο αποτελείται από ένα εκατομμύρια περίπου όμοια αντίγραφα.



- Το κάθε μικροσφαιρίδιο τοποθετείται στο εσωτερικό ενός πηγαδιού μεγέθους όσο και του μικροσφαιριδίου.
- Στη συνέχεια ξεκινά ο κύκλος της πυραλληλούχησης.
- Το σήμα καταγράφεται από CCD κάμερες.

Sequencing

7.5 hours

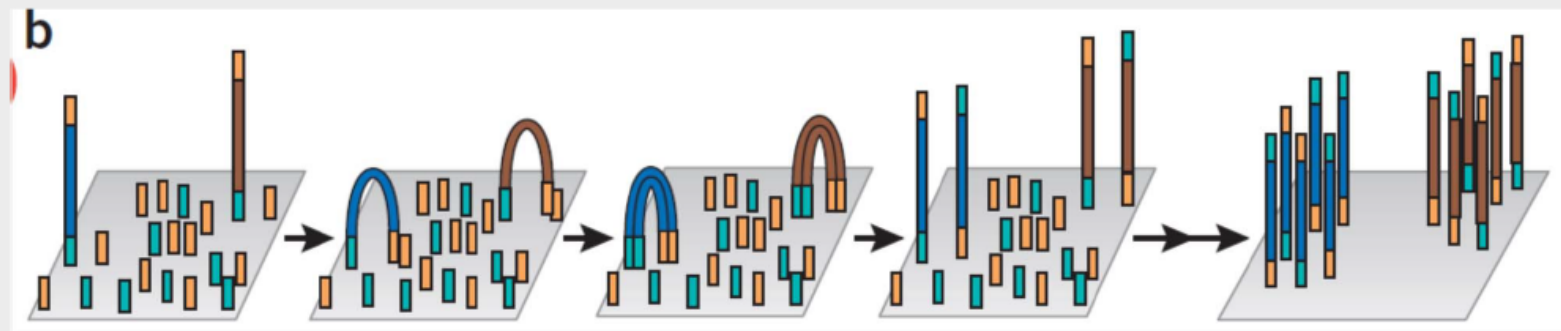


- Well diameter: average of 44 μm
- 400,000 reads obtained in parallel
- A single cloned amplified sstDNA bead is deposited per well

Amplified sstDNA library beads

Quality filtered bases

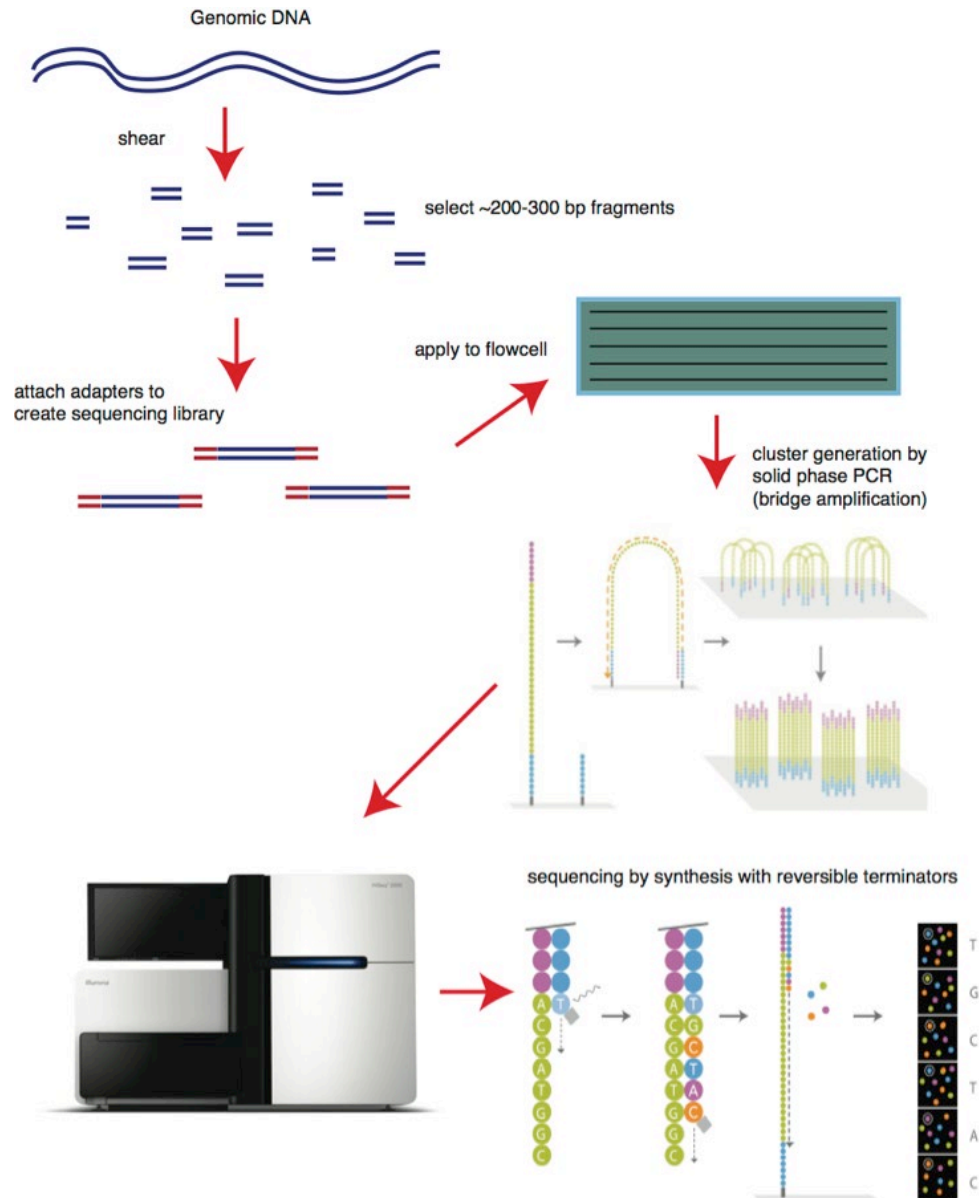
Bridge PCR



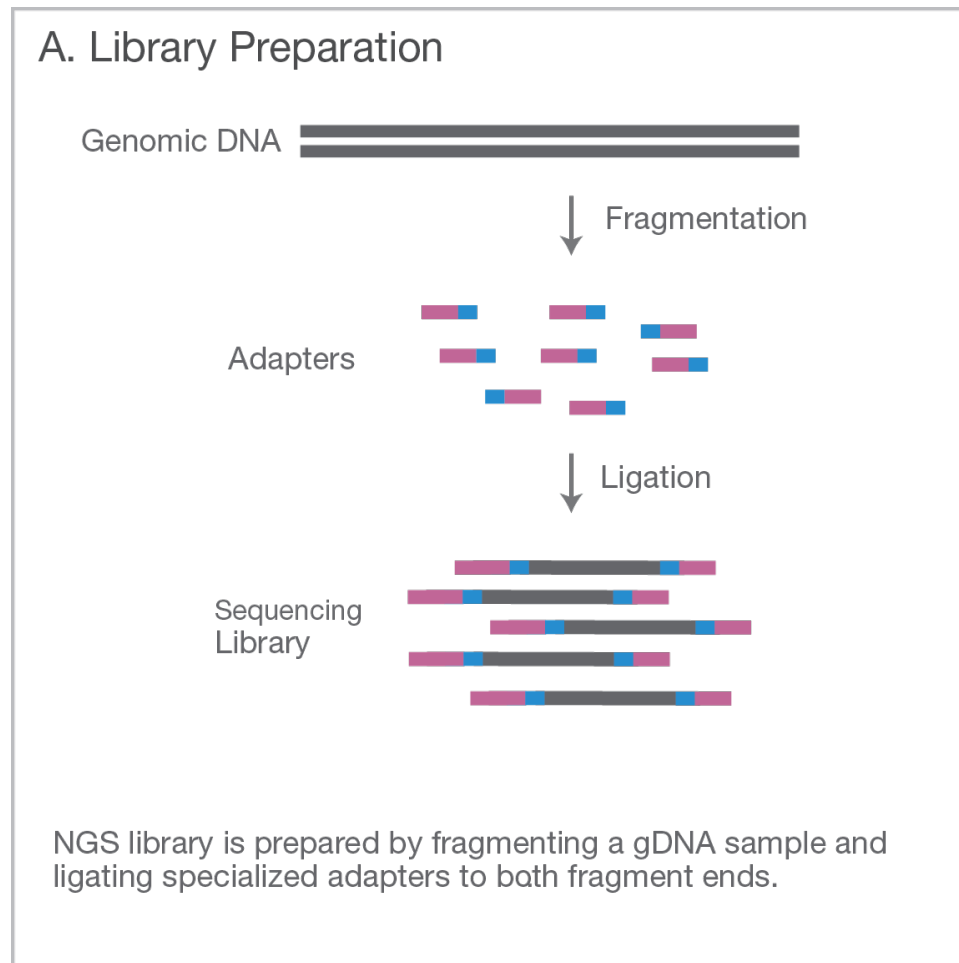
- DNA fragments are flanked with adaptors
- A flat surface coated with two types of primers, corresponding to the adaptors
- Amplification proceeds in cycles, with one end of each bridge tethered to the surface
- Used by illumina

ILLUMINA sequencing systems

- Sequencing by synthesis
- Προετοιμασία δείγματος: κλασμάτωση, adaptor ligation, ακινητοποίηση
- In situ bridge amplification
- Αντίδραση με RT dNTPs σημασμένα με διαφορετικό φθοριόχρωμα το καθένα
- Ακολουθία από snap shots θα δώσει την αλληλουχία των νουκλεοτιδίων που ενσωματώθηκαν



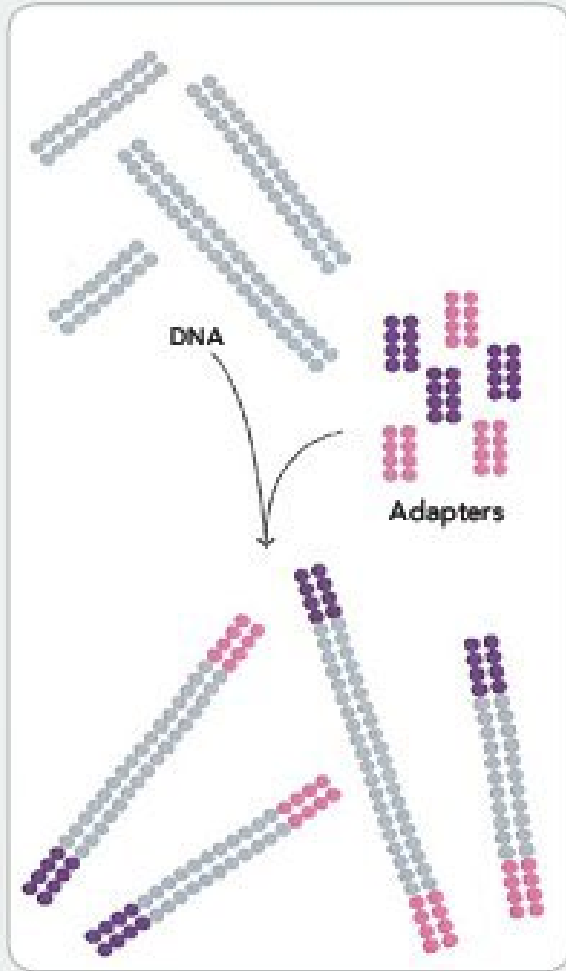
Η ροή εργασίας περιλαμβάνει τέσσερα βασικά βήματα:



Προετοιμασία της βιβλιοθήκης: Η βιβλιοθήκη κατασκευάζεται με τυχαίο κατακερματισμό του δείγματος DNA ή cDNA, και ακολουθεί σύνδεση προσαρμογέων στο 5' και 3'

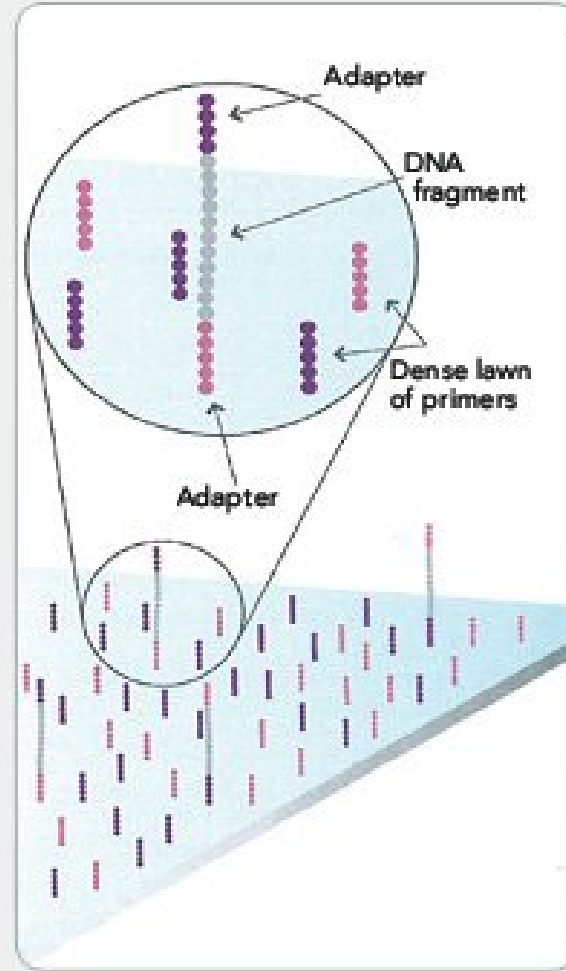
Illumina – sequencing by synthesis (Solexa)

1. PREPARE GENOMIC DNA SAMPLE



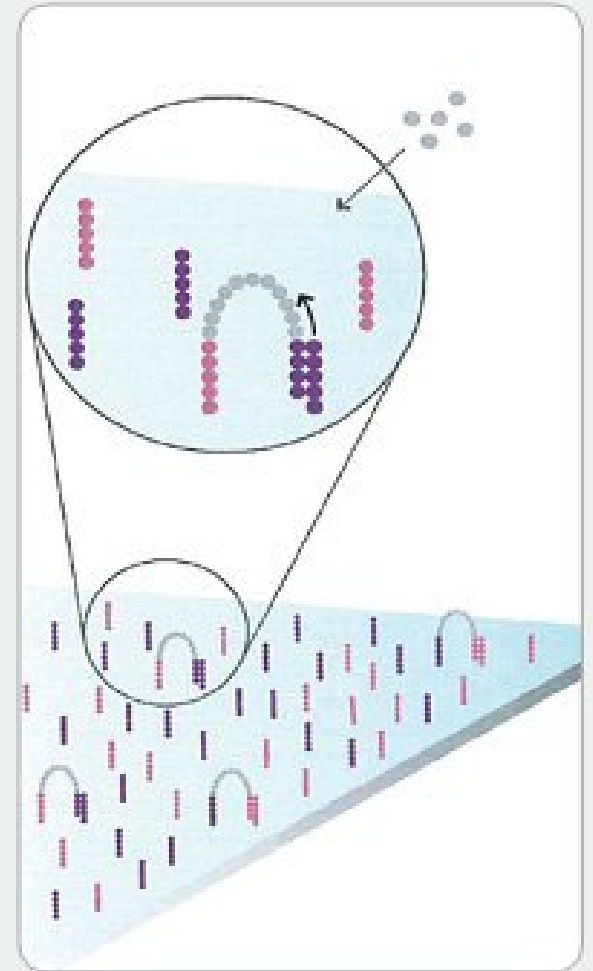
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

2. ATTACH DNA TO SURFACE



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

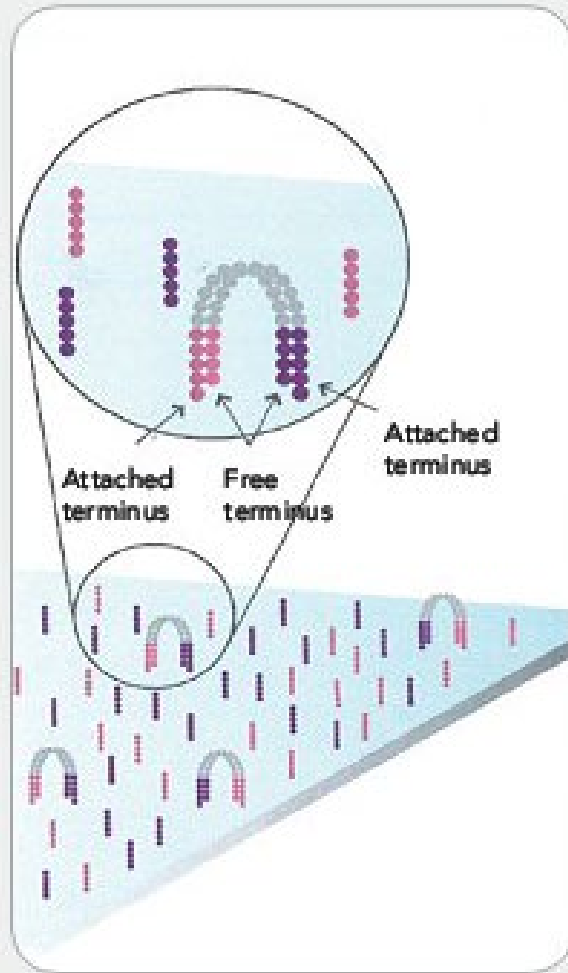
3. BRIDGE AMPLIFICATION



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

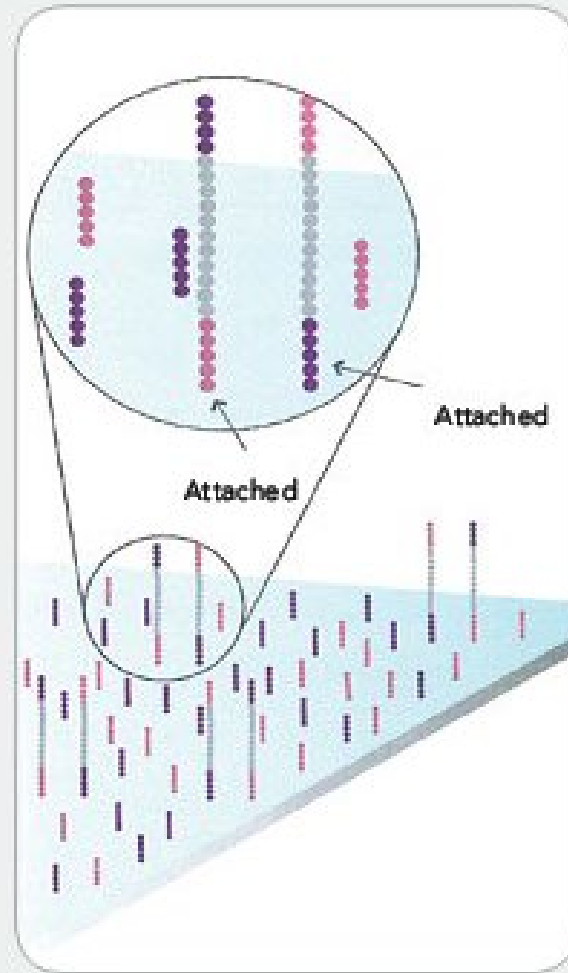
Illumina – sequencing by synthesis (Solexa)

4. FRAGMENTS BECOME DOUBLE STRANDED



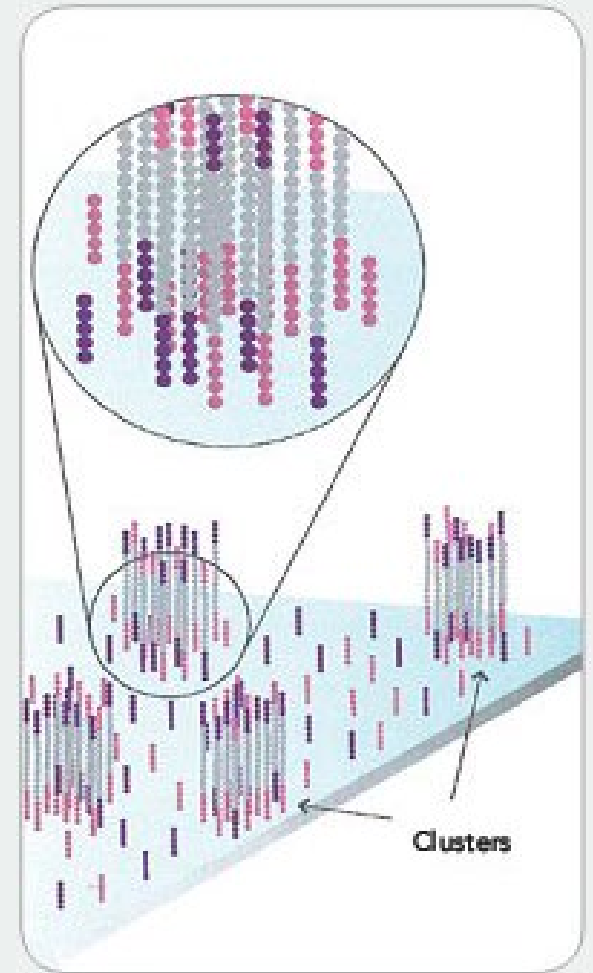
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

5. DENATURE THE DOUBLE-STRANDED MOLECULES



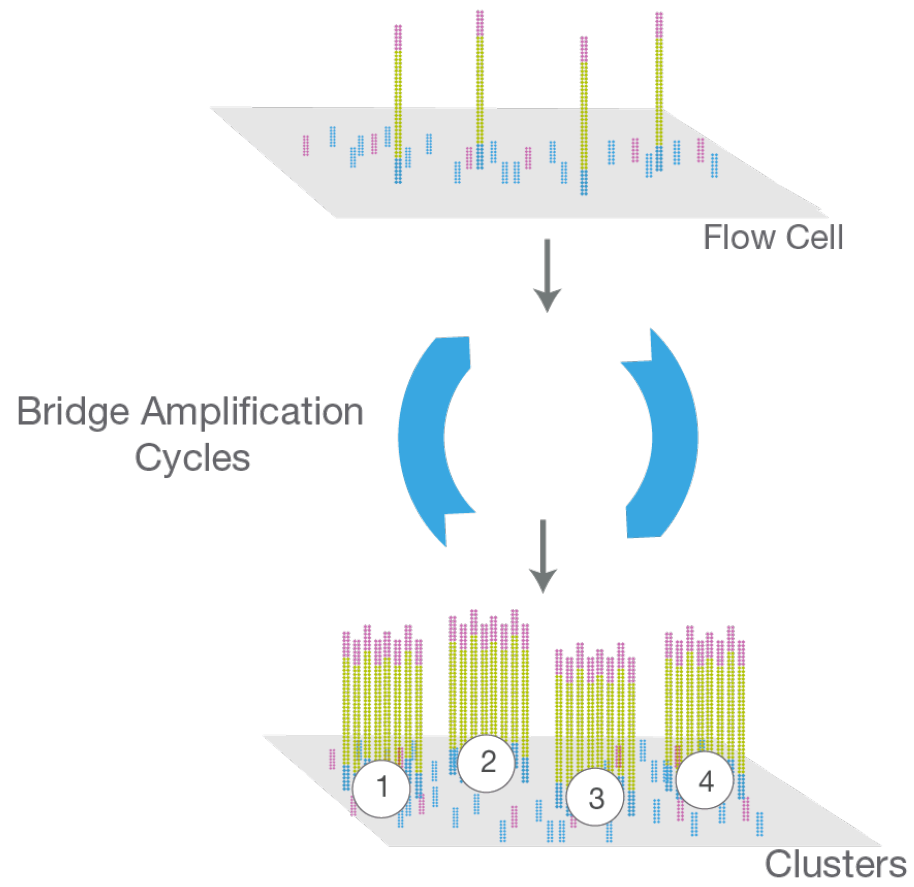
Denaturation leaves single-stranded templates anchored to the substrate.

6. COMPLETE AMPLIFICATION



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

B. Cluster Amplification



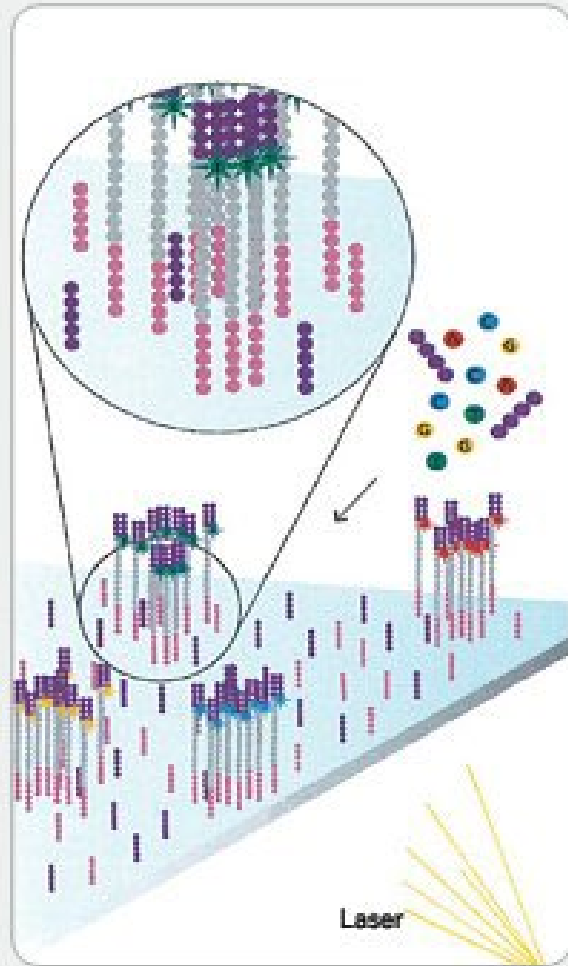
Library is loaded into a flow cell and the fragments are hybridized to the flow cell surface. Each bound fragment is amplified into a clonal cluster through bridge amplification.

Δημιουργία συστάδων - clusters:

- Για τη δημιουργία συστάδων, η βιβλιοθήκη φορτώνεται σε μια κυψέλη ροής όπου τα θραύσματα δεσμεύονται σε ολιγονουκλεοτίδια που είναι συμπληρωματικά προς τους προσαρμογείς βιβλιοθήκης και δεσμευμένα σε μια επιφάνεια.
- Κάθε θραύσμα στη συνέχεια ενισχύεται σε διακριτές, κλωνικές ομάδες μέσω bridge amplification έτσι ώστε να σχηματιστούν συστάδες θραυσμάτων DNA από κάθε θραύσμα.
- Αποτέλεσμα της παραπάνω διαδικασίας είναι ότι κάθε βιβλιοθήκη θραυσμάτων αποτελείται πλέον από εκατοντάδες εκατομμύρια μοναδικά συμπλέγματα (clusters).
- Όταν ολοκληρωθεί η δημιουργία συμπλέγματος, τα templates είναι έτοιμα για αλληλούχηση.

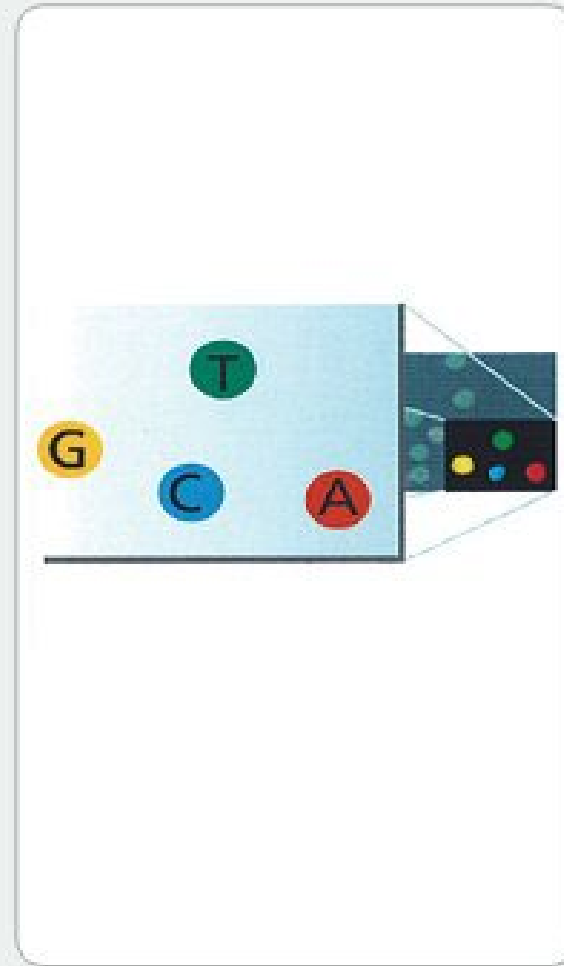
Illumina – sequencing by synthesis (Solexa)

7. DETERMINE FIRST BASE



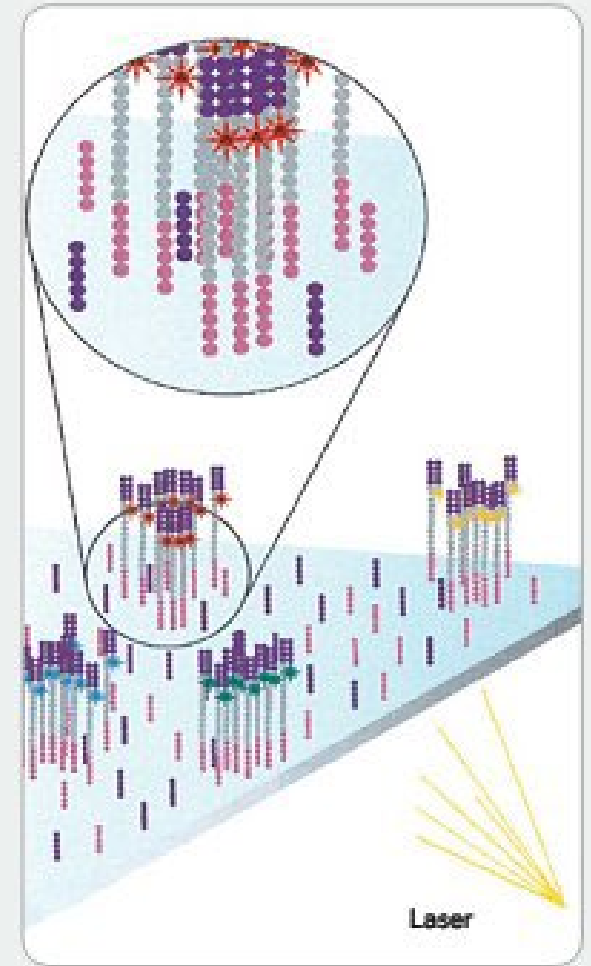
First chemistry cycle: to initiate the first sequencing cycle, add all four labeled reversible terminators, primers and DNA polymerase enzyme to the flow cell.

8. IMAGE FIRST BASE



After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

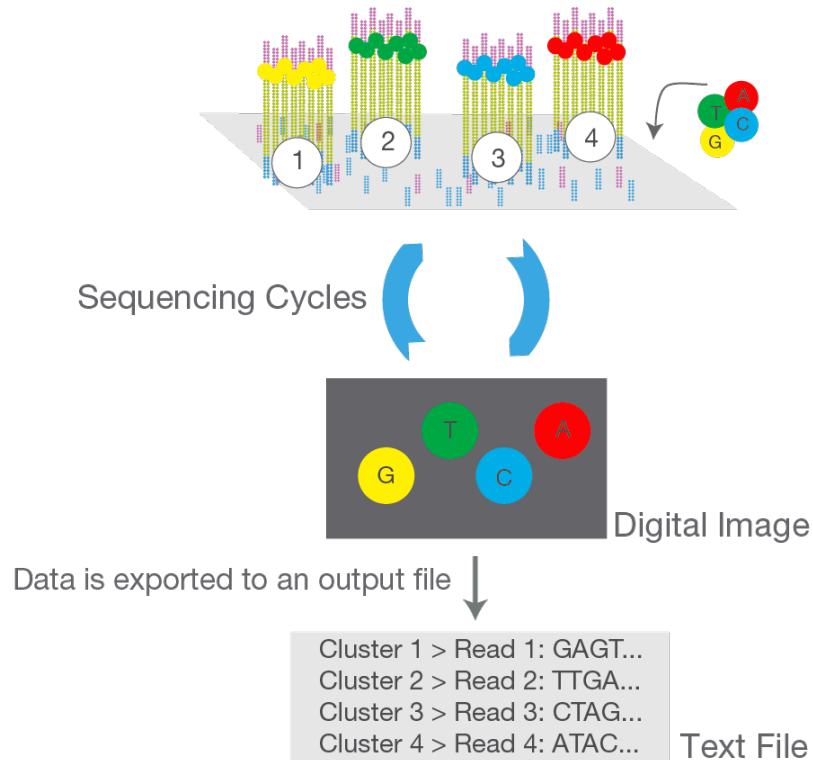
9. DETERMINE SECOND BASE



Second chemistry cycle: to initiate the next sequencing cycle, add all four labeled reversible terminators and enzyme to the flow cell.

Αλληλούχηση

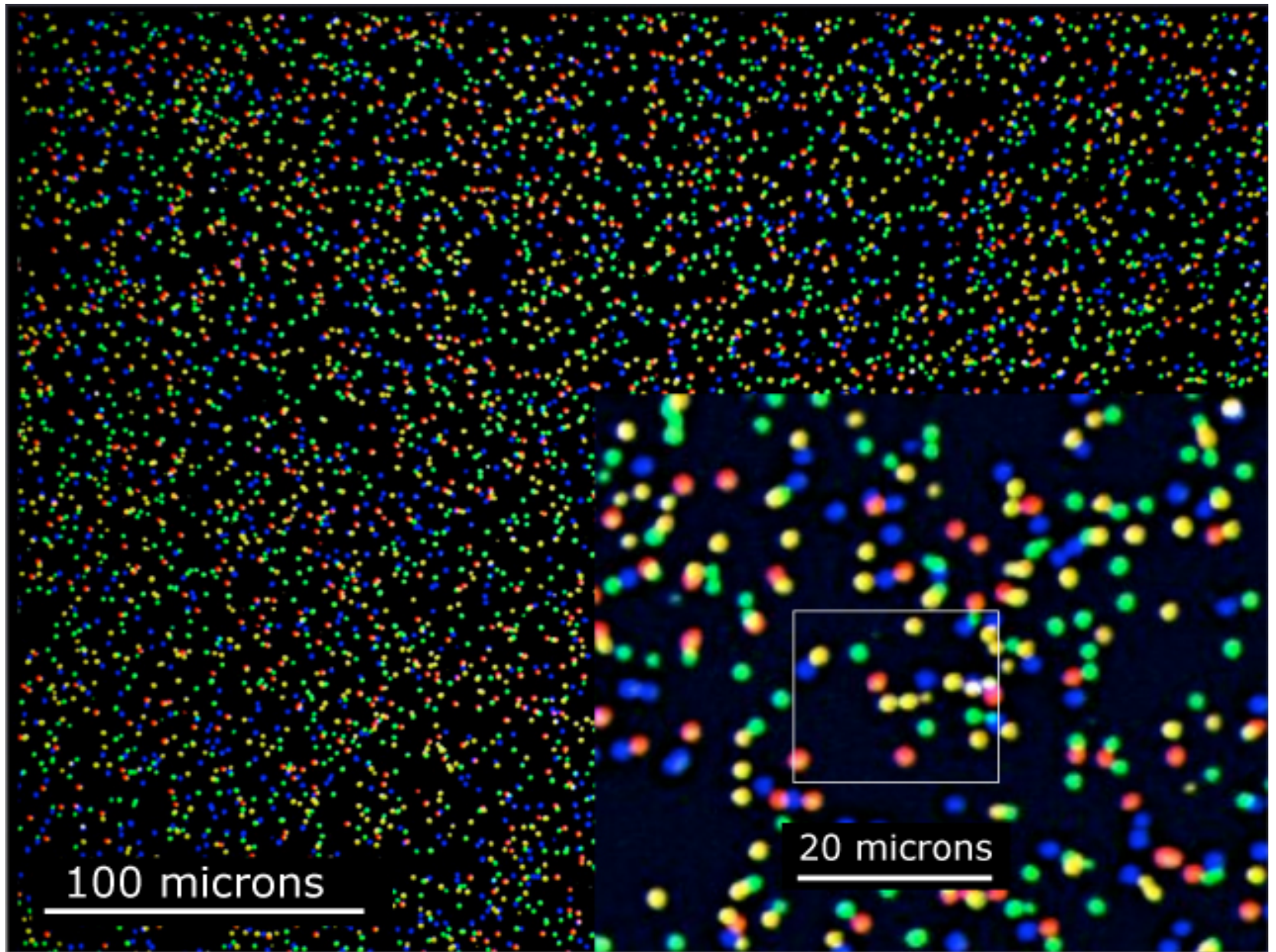
C. Sequencing



Sequencing reagents, including fluorescently labeled nucleotides, are added and the first base is incorporated. The flow cell is imaged and the emission from each cluster is recorded. The emission wavelength and intensity are used to identify the base. This cycle is repeated “n” times to create a read length of “n” bases.

- Η τεχνολογία Illumina χρησιμοποιεί μια μέθοδο αναστρέψιμου τερματισμού που ανιχνεύει μεμονωμένες βάσεις καθώς ενσωματώνονται σε κλώνους μήτρας DNA των συστάδων
- Το μίγμα που χρειάζεται για την αντίδραση της αλληλούχησης και σύνθεσης DNA παρέχεται στην επιφάνεια της κυψέλης ροής και περιέχει τέσσερα αναστρέψιμα νουκλεοτίδια τερματισμού το κάθε ένα από τα οποία είναι σημασμένο με διαφορετική φθορίζουσα χρωστική και το 3'-OH άκρο τους μπλοκαρισμένο χημικά έτσι ώστε κάθε ενσωμάτωση να είναι ένα μοναδικό γεγονός.
- Οι τέσσερις φθορίζουσες με τις βάσεις πλησιάζουν την βάση του cluster αλλά μόνο μία θα ενωθεί. Και οι τέσσερις βάσεις ανταγωνίζονται μεταξύ τους για να συνδεθούν με το εκμαγείο. Αυτός ο ανταγωνισμός εξασφαλίζει την υψηλότερη δυνατή ακρίβεια. Μόλις το λέιζερ – CCD camera ανιχνεύσει τη συμπληρωματική βάση (από το χρώμα που εκπέμπει), η φθορίζουσα χρωστική αφαιρείται και μένει η βάση μετά από ξέπλυμα. Το ίδιο γίνεται και για την επόμενη βάση της αλυσίδας του cluster μέχρι να τερματιστεί. Έτσι δημιουργούνται συμπληρωματικές αλυσίδες των clusters.

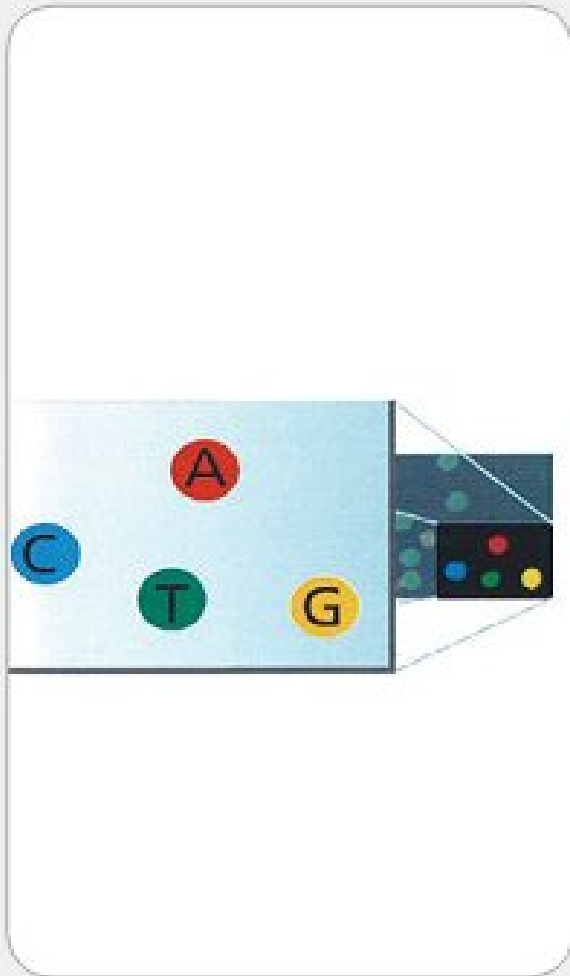
ILLUMINA sequencing systems



Illumina instant shot

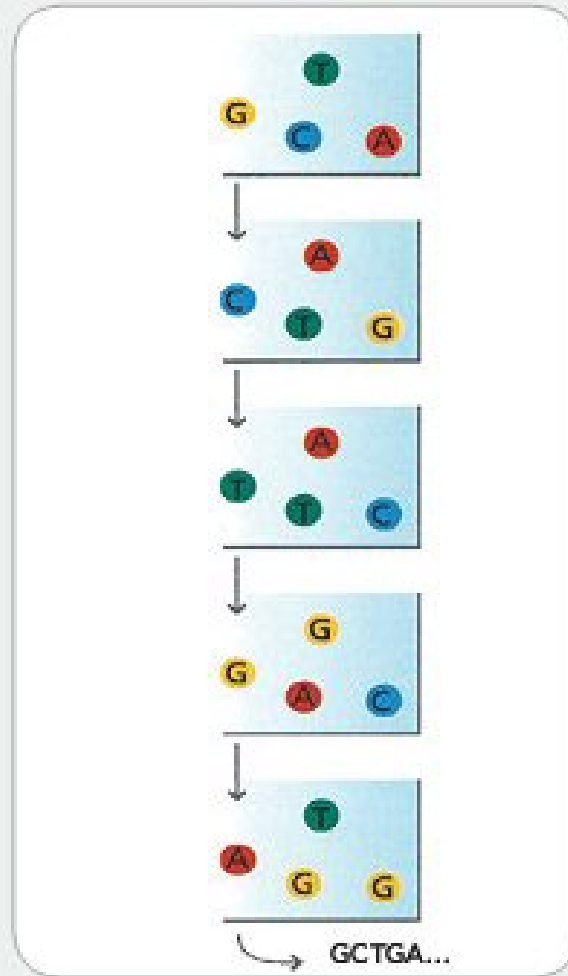
Illumina – sequencing by synthesis (Solexa)

10. IMAGE SECOND CHEMISTRY CYCLE



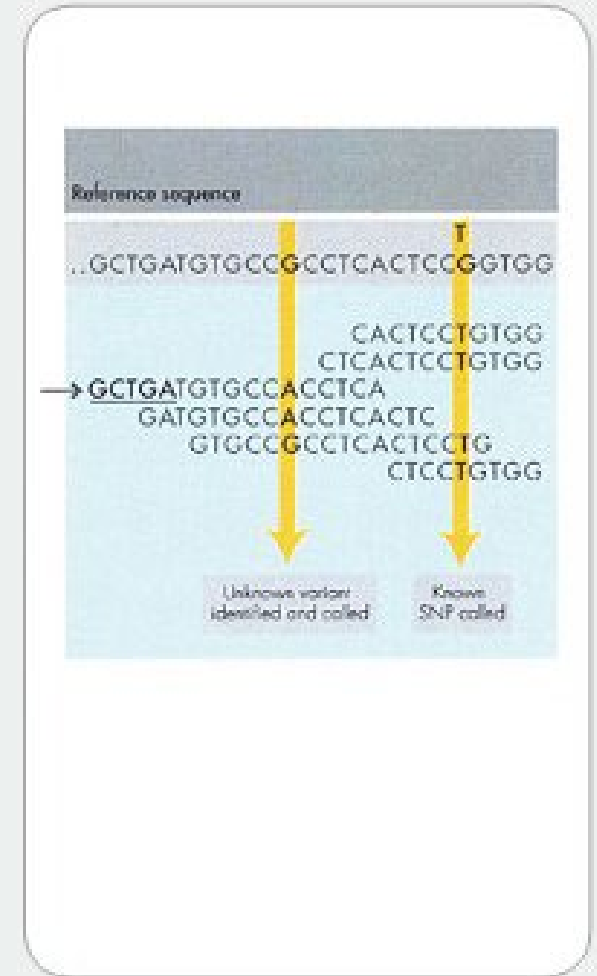
After laser excitation, collect the image data as before. Record the identity of the second base for each cluster.

11. SEQUENCE READS OVER MULTIPLE CHEMISTRY CYCLES



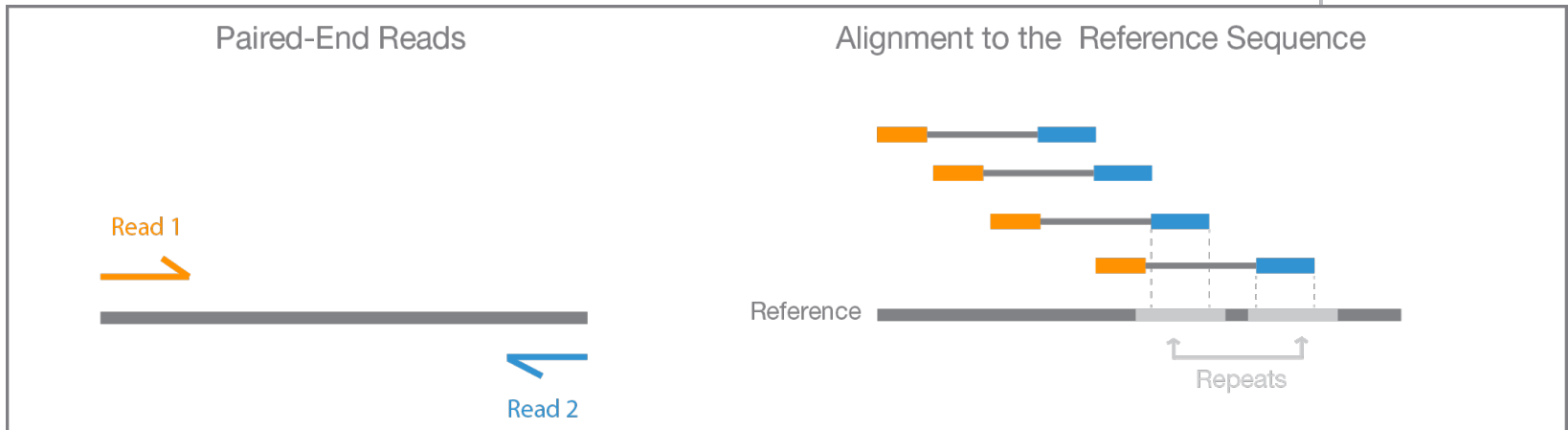
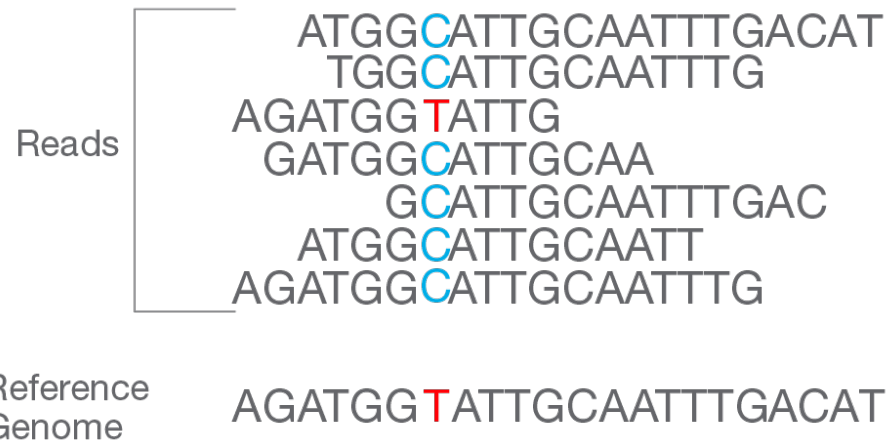
Repeat cycles of sequencing to determine the sequence of bases in a given fragment a single base at time.

12. ALIGN DATA



Align data, compare to a reference, and identify sequence differences.

D. Alignment and Data Analysis



Ion Torrent, Life Technologies

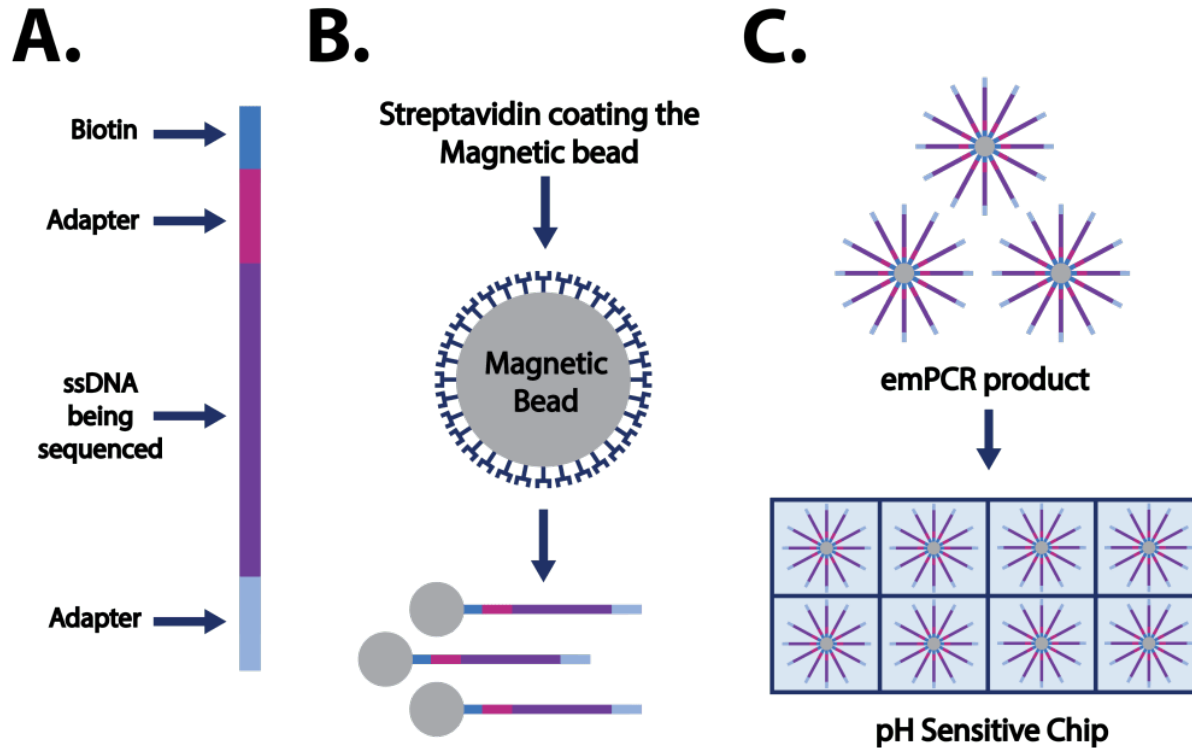


- Sequencing by synthesis
- Χρησιμοποιεί semiconductor technology
- Δεν χρησιμοποιεί Terminator dNTPs ούτε σήμανση
- Εισάγεται ένα νουκλεοτίδιο τη φορά
- Ανιχνεύει τα H^+ που απελευθερώνονται με την ενσωμάτωση ενός νουκλεοτιδίου στην αυξανόμενη αλυσίδα
- Η απελευθέρωση H^+ μετριέται ως αλλαγή στο pH ή στην αγωγιμότητα
- Μπορεί να διαβάζει μεγαλύτερα κομμάτια από τους ανταγωνιστές του
- Γρήγορο και οικονομικό
- Πιο ικανό??? στο «διάβασμα» ομοπολυμερών

Ion Torrent-Life tech-ABI

- Χρειάζεται ΜΟΝΟ ένα μικροτσιπ ιόντων για να διαβάσει ένα βακτηριακό γονιδίωμα (25 Mb) σε 2 ώρες
- Υπάρχουν 1,2 εκατομμύρια θέσεις για αντιδράσεις-(ανάλογες με 454)
- Βασίζεται στην ικανότητα του μηχανήματος να αναγνωρίζει την αλλαγή του pH από τα ιόντα που απελευθερώνονται κατά τη προσθήκη των βάσεων κατά τη σύνθεση του DNA σε κάθε θέση αντίδρασης
- Μέσο μέγεθος κομματιών 100bp

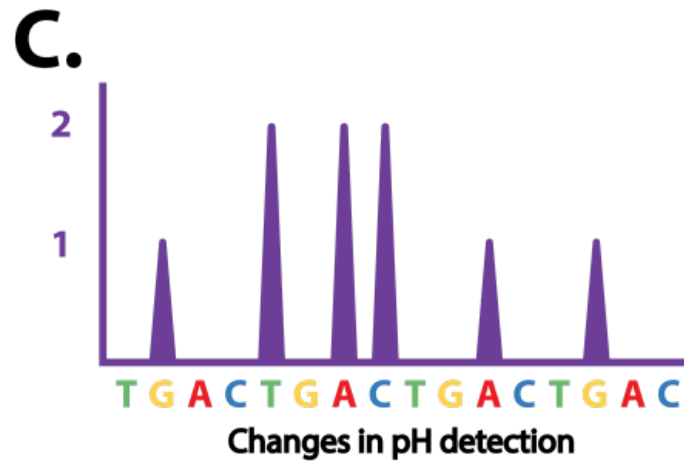
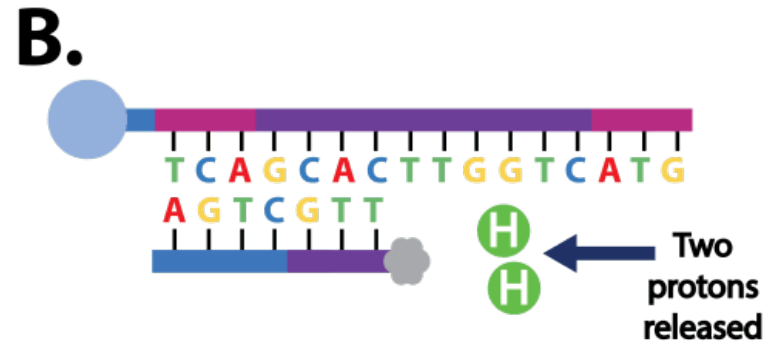
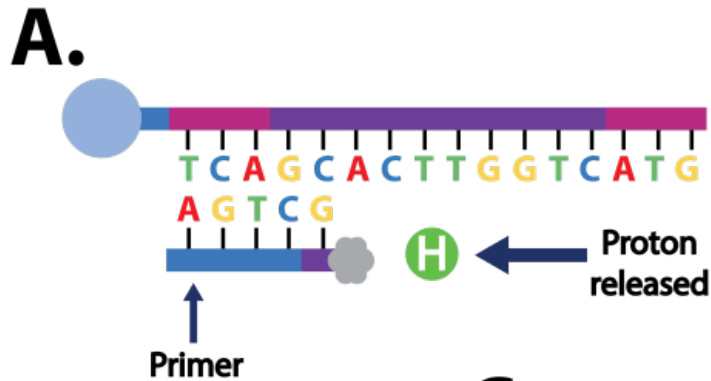
ION TORRENT DNA LIBRARY PREPARATION



<https://apollo-institute.org/ion-torrent-sequencing/>

- Το μόριο DNA στόχος διασπάται σε περίπου 400 bp μονόκλιωνα τμήματα DNA (ssDNA) και δύο διαφορετικούς προσαρμογείς, με βιοτίνη, οι οποίοι στη συνέχεια προστίθενται στα άκρα.
- Η βιβλιοθήκη ssDNA που προκύπτει συνδέεται μεμονωμένα (ανά μόριο DNA) σε ένα μαγνητικό σφαιρίδιο μέσω της αλληλεπίδρασης στρεπταβιδίνης και βιοτίνης, και ενισχύεται μέσω αλυσιδωτής αντίδρασης πολυμεράσης γαλακτώματος (emPCR).
- Ως αποτέλεσμα, εκατομμύρια από τα ίδια θραύσματα ssDNA συνδέονται σε ένα μόνο σφαιρίδιο. Στη συνέχεια, κάθε σφαιρίδιο τοποθετείται σε ένα τσιπ ευαίσθητο στο pH με χιλιάδες πηγαδακία.

ION TORRENT SEQUENCING & DATA ANALYSIS



- Για να ξεκινήσει η αλληλούχηση, ένα από τα τέσσερα dNTP προστίθεται και ξεπλένεται ανάλογα με το αν ανιχνεύεται σήμα.
- Κάθε φορά που ένα dNTP ενσωματώνεται επιτυχώς στον επιμηκυνόμενο κλώνο από το DNAP, απελευθερώνονται ιόντα υδρογόνου με αποτέλεσμα τη μείωση του pH που ανιχνεύεται από το τσιπ που είναι ευαίσθητο στο pH. Μια μεγαλύτερη μείωση στο pH υποδηλώνει ότι ενσωματώνονται διαδοχικά dNTPs.
- Τέλος, τα δεδομένα αναλύονται από υπολογιστή προκειμένου να ληφθεί η αλληλουχία DNA του αρχικού μορίου DNA.
- Αν και η αλληλουχία ιόντων torrent προσφέρει μια άλλη τεχνολογία ταχείας αλληλουχίας για χρήση από τους επιστήμονες, όπως ακριβώς και με το pyrosequencing, ένα σημαντικό μειονέκτημα είναι ότι το σήμα γίνεται λιγότερο καθαρό όταν αλληλουχούνται επαναλαμβανόμενες περιοχές ομοπολυμερούς

Single-molecule, real-time DNA - SMRT sequencing

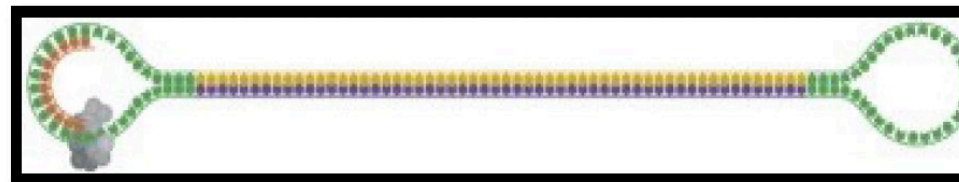
- Αλληλούχηση 3^{ης} γενιάς
- Η μέθοδος βασίζεται στο φθορισμό
- Δεν είναι απαραίτητη η ενίσχυση με αλυσιδωτή αντίδραση πολυμεράσης PCR πριν τον προσδιορισμό αλληλουχίας
- Το παραγόμενο σήμα συλλαμβάνεται σε πραγματικό χρόνο καθώς παρακολουθείται η πορεία της ενζυματικής αντίδρασης.
- Χρησιμοποιούνται τέσσερα φώσφορο-συνδεδεμένα νουκλεοτίδια διαφορετικών χρωμάτων
- Ο φθορισμός απελευθερώνεται ταυτόχρονα με την νουκλεοτιδική ενσωμάτωση

Pacific Biosciences

- ❖ Βασίζεται στον προσδιορισμό αλληλουχίας με σύνθεση και επιτρέπει την παρατήρηση της σύνθεσης DNA την ώρα που συμβαίνει, σε πραγματικό χρόνο.
- ❖ Αυτό μπορεί να πραγματοποιηθεί με την χρήση (ZMW) οι οποίοι είναι ειδικοί θάλαμοι νανοφωτονικής απεικόνισης.
- ❖ Σε κάθε SMRT cell κατασκευάζονται περίπου 75,000 ZMW με τους οποίους μπορεί να γίνει ανίχνευση 75,000 αντιδράσεων αλληλούχισης παράλληλα.
- ❖ Για την προετοιμασία της βιβλιοθήκης τα θραύσματα DNA ενώνονται με προσαρμογείς βρόγχου - φουρκέτας και στις δύο πλευρές.



Εικόνα 1.15 SMRT cell

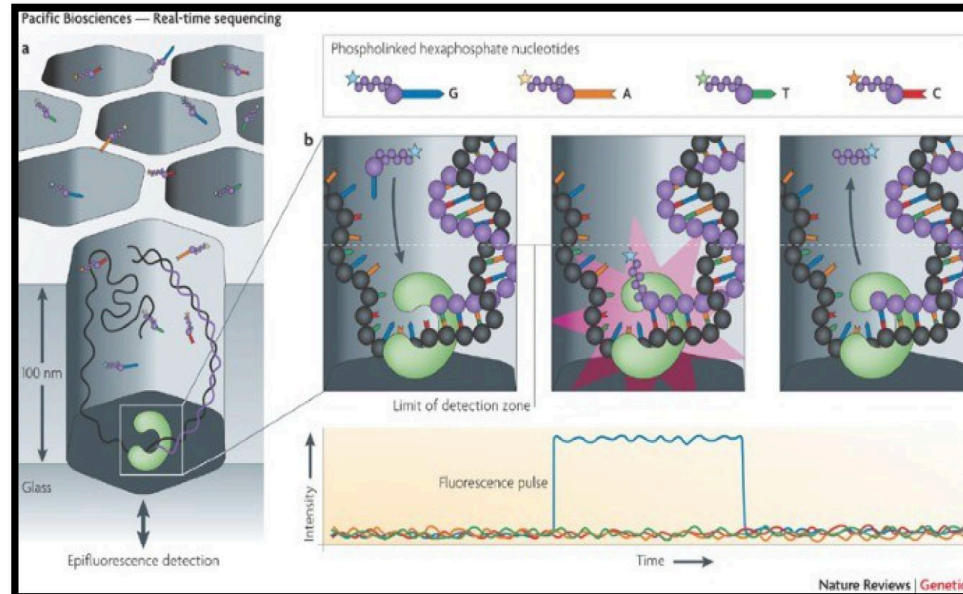


Εικόνα 1.16 Πρότυπο SMRT bell

Pacific Biosciences

Στο ZMW είναι προσαρτημένο ένα μόριο πολυμεράσης μέσω του συστήματος βιοτίνης - στρεπταβιδίνης.

Ο βρόγχος DNA κινείται μέσω της πολυμεράσης και κάθε φορά που ένα νουκλεοτίδιο συγκρατείται ένας φωσφορικός δεσμός σπάει και παράγεται ένας παλμός φωτός ο οποίος καταγράφεται από μια κάμερα CCD και ερμηνεύεται.



Εικόνα 1.17 Αλληλούχιση πραγματικού χρόνου με SMRT. Τα σημασμένα νουκλεοτίδια εισέρχονται μέσα στα ZMW, φθάνουν στην πολυμεράση και διαχέονται ξανά. Όταν γίνεται σωστή ενσωμάτωση νουκλεοτιδίου υπάρχει υψηλότερη ένταση σήματος η οποία καταγράφεται από κάμερα CCD

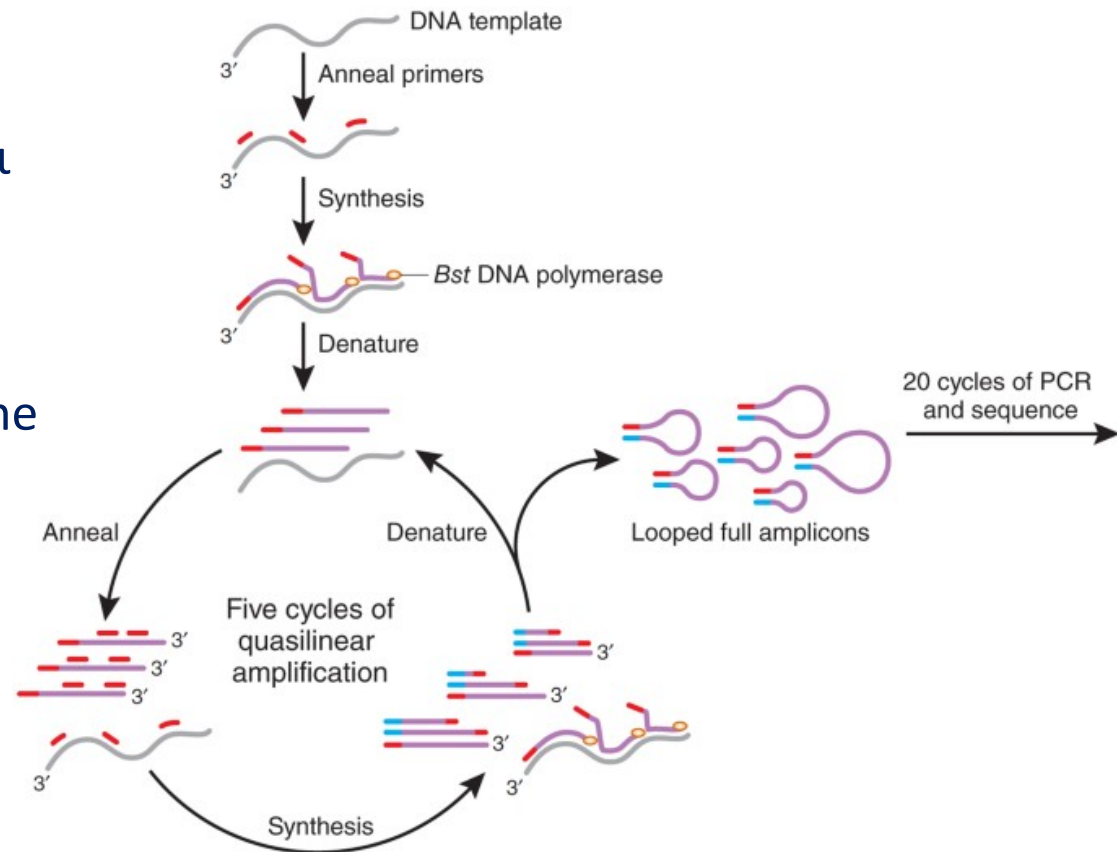
Το γονιδίωμα ενός κυττάρου

Οι νέες τεχνολογίες έχουν επιτρέψει την αλληλούχηση του γονιδιώματος ενός μόνο κυττάρου

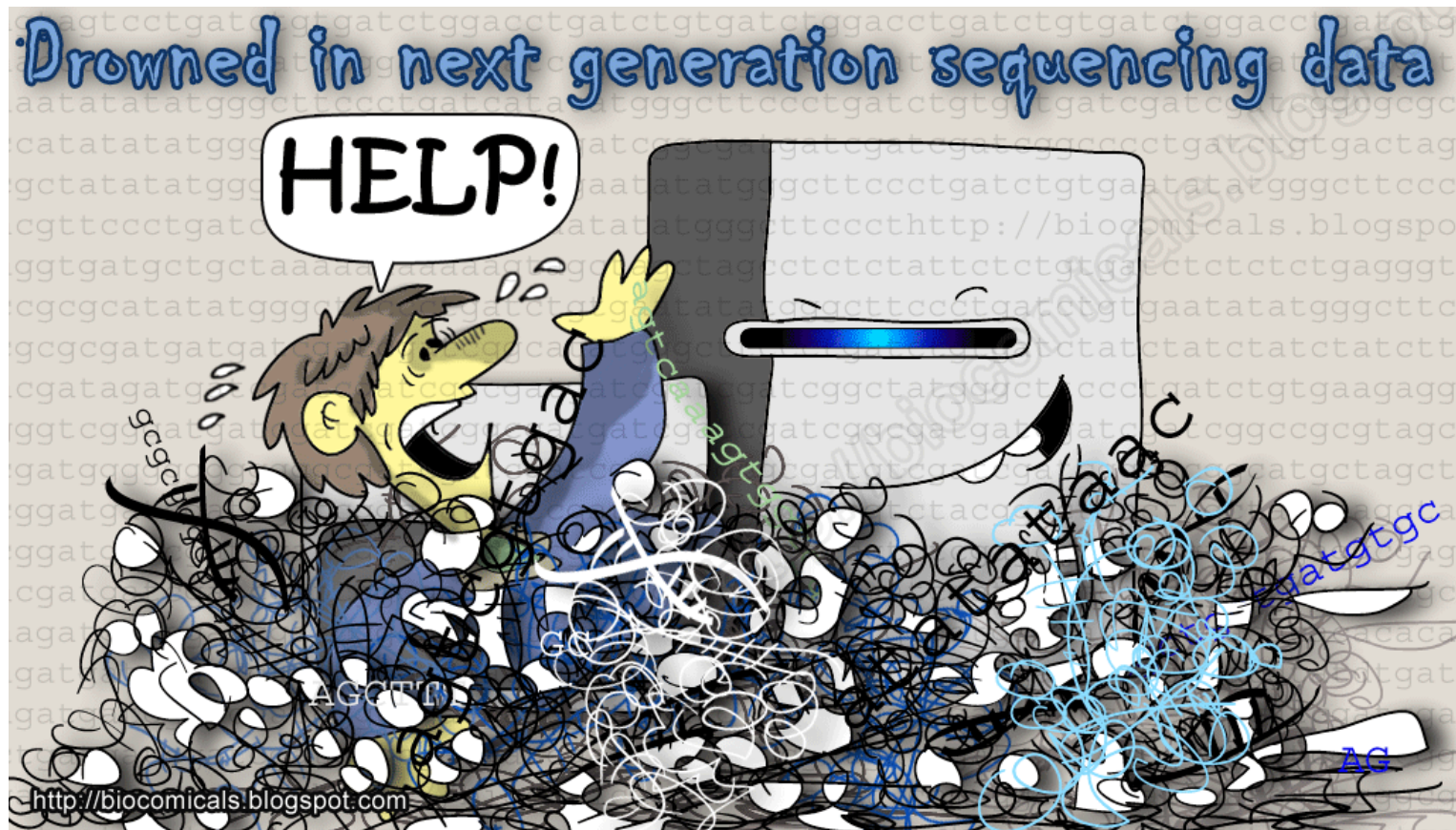
Μεθοδολογία που ακολουθείται για την αλληλούχηση από ένα μόνο κύτταρο

Single-cell sequencing in its prime

Lasken R.S., 2013, Nature Biotechnology 31, 211–212



Το NextGen sequencing έχει ως αποτέλεσμα τη δημιουργία εκατοντάδων χιλιάδων αλληλουχιών οι οποίες πρέπει να επεξεργαστούν με εργαλεία βιοπληροφορικής για την εξαγωγή της ολοκληρωμένης αλληλουχίας.



Η εποχή των -omics

Phenome: πλήρης περιγραφή του φαινότυπου (χρήση νοκ-αουτ γονιδίων)

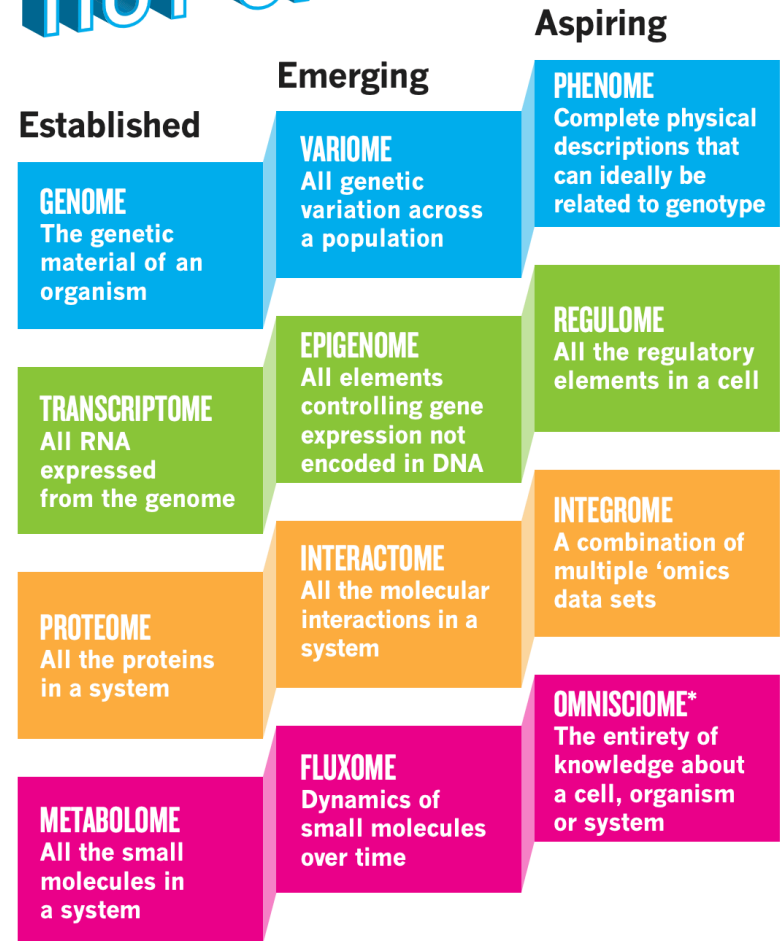
Interactome: όλες οι μοριακές αλληλεπιδράσεις εντός κυττάρου

Intergrome: η ενσωμάτωση των δεδομένων σε βάσεις δεδομένων και η επικοινωνία τους

Incidentalome: τα γενετικά δεδομένα που ανακαλύπτονται «κατά λάθος» π.χ. 99 κοινές γενετικές παραλλαγές σχετιζόμενες με ασθένειες

Toxome: πως αντιδρά ο οργανισμός σε τοξικές ενώσεις

HOT OR NOT



*Nature's proposed addition to the scientific nomenclature.