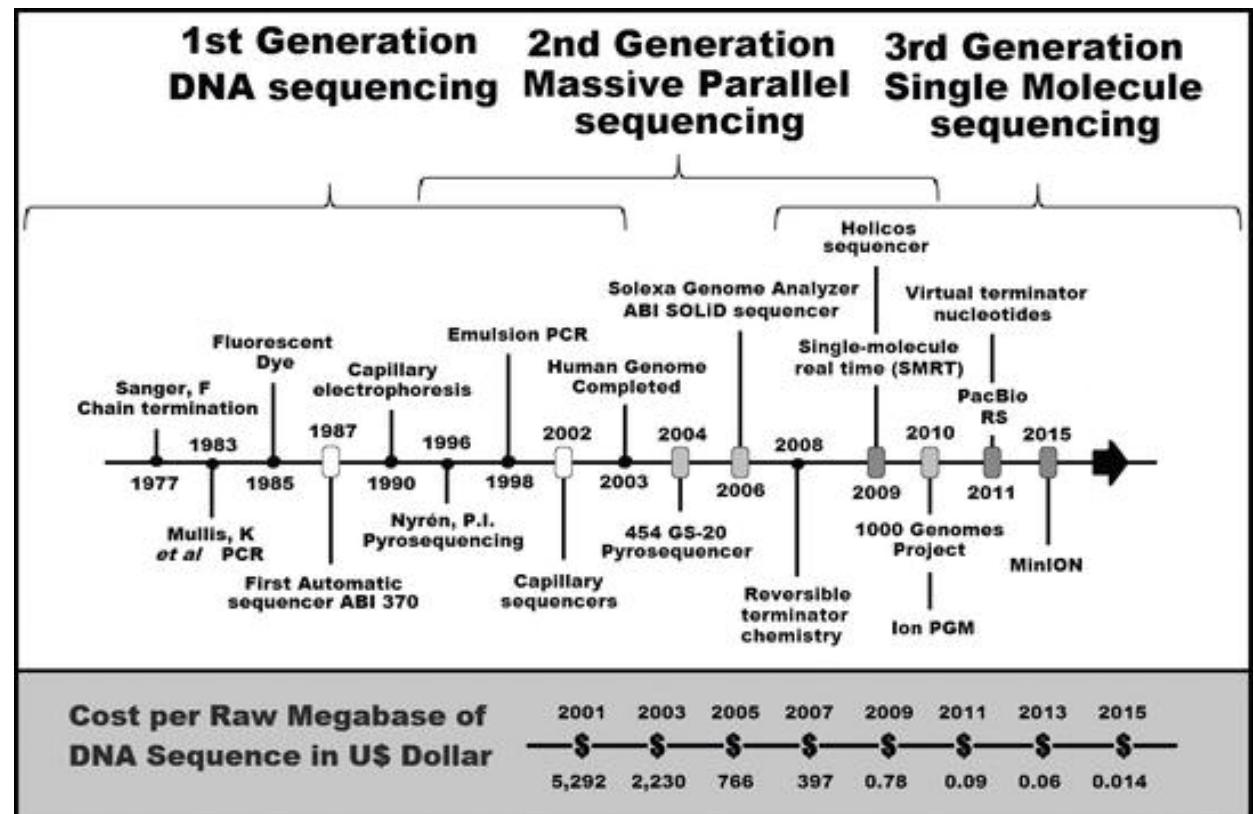


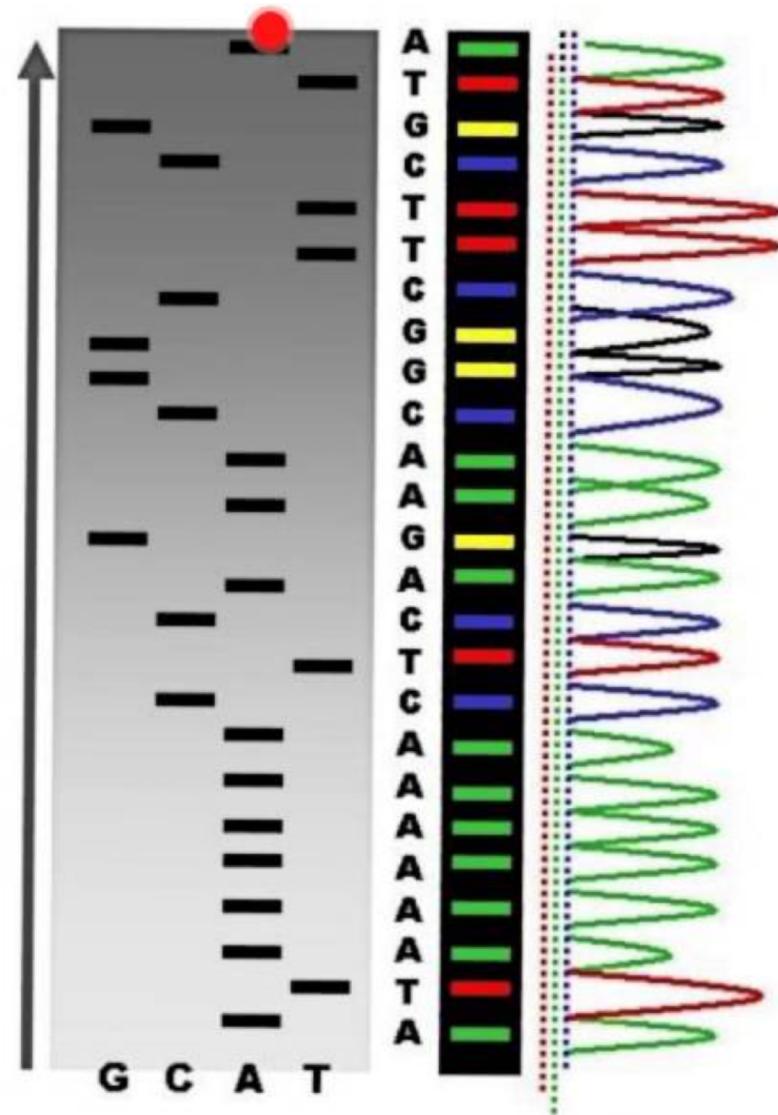
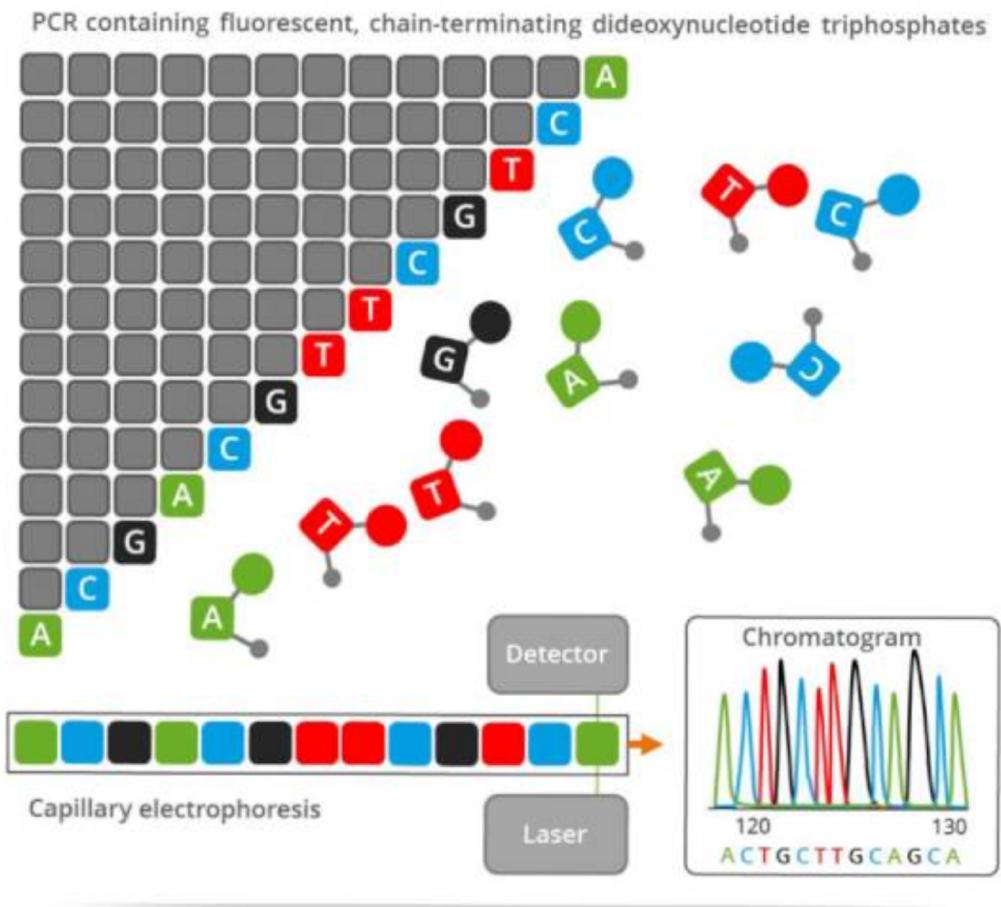
Dizileme Teknolojileri

Dizileme teknolojileri

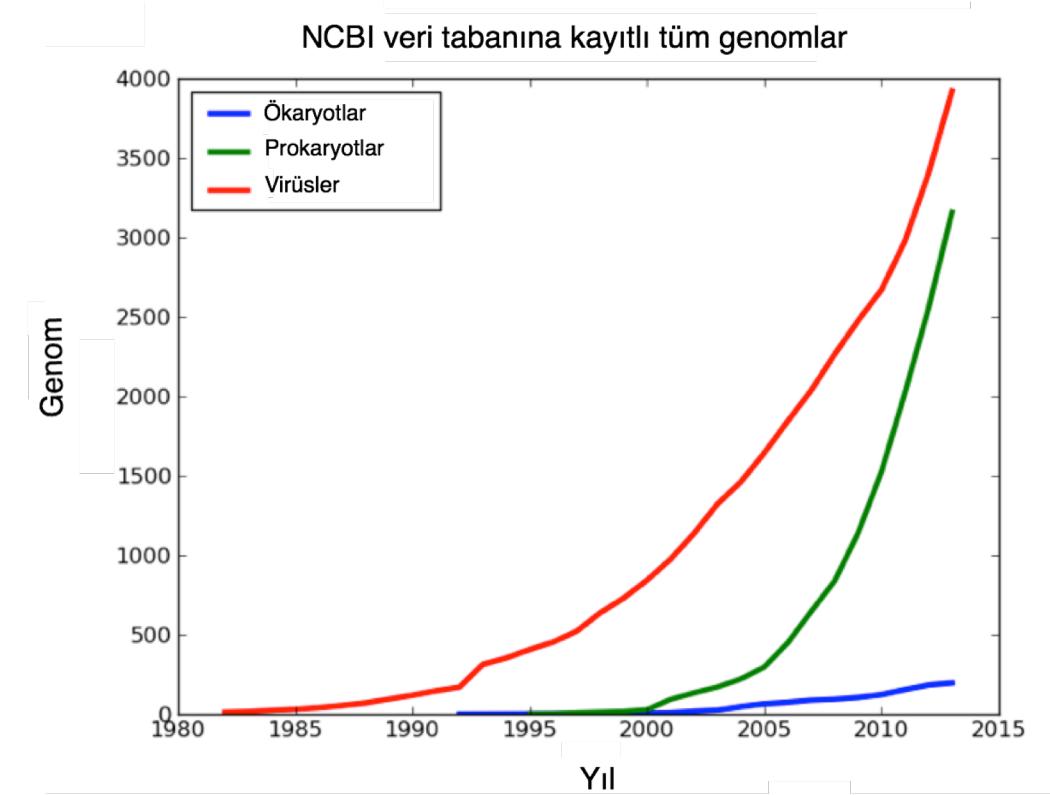
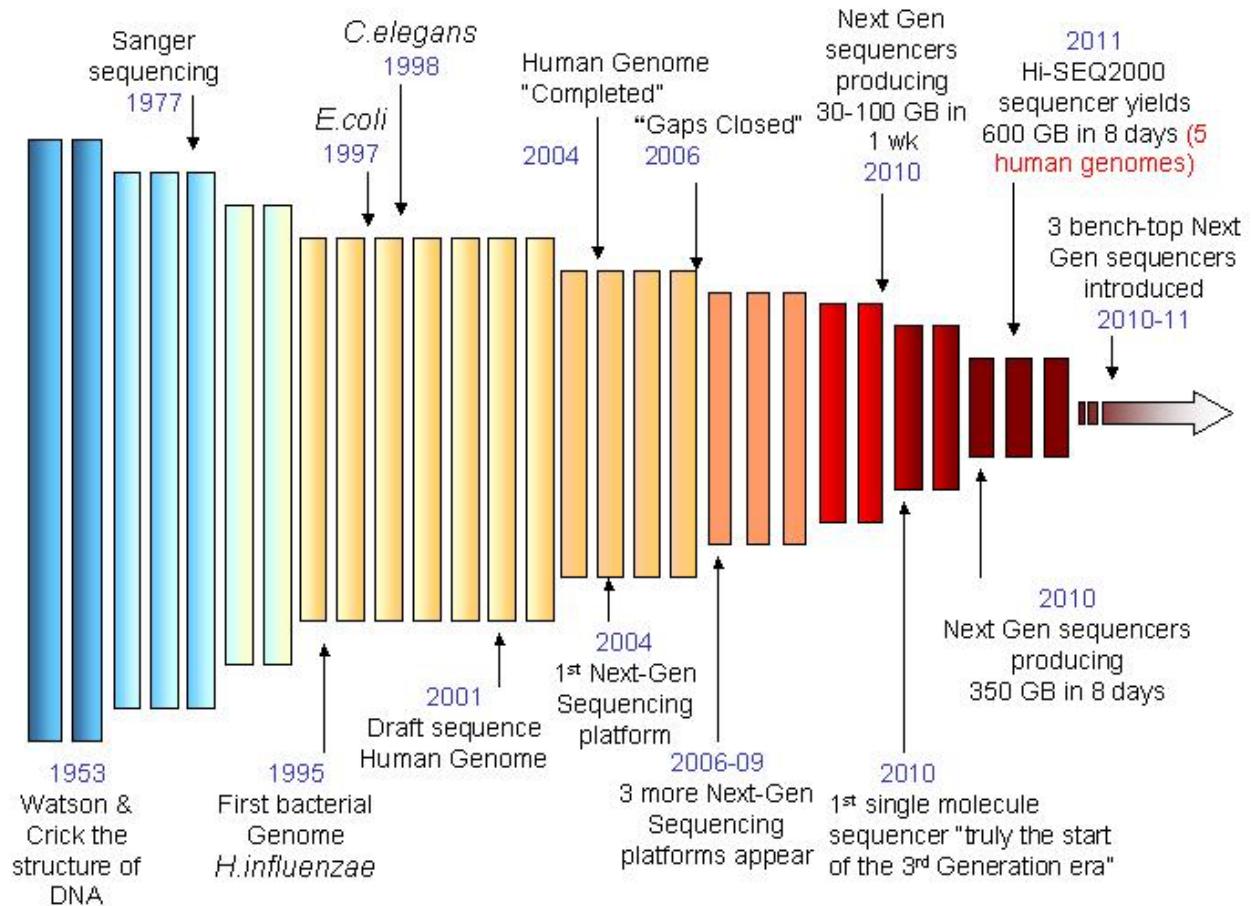
- Sanger dizileme (1977`den beri)
- Yeni nesil dizileme (2005`den beri)
 - * 454
 - * Ion Torrent
 - * Illumina
- 3. nesil dizileme (2012`den beri)
 - * PacBio ya da SMRT dizileme
 - * Oxford Nanopore



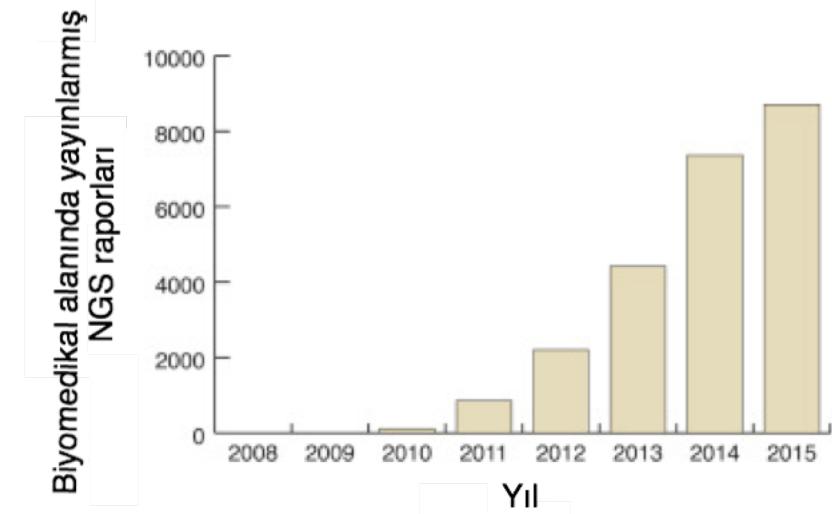
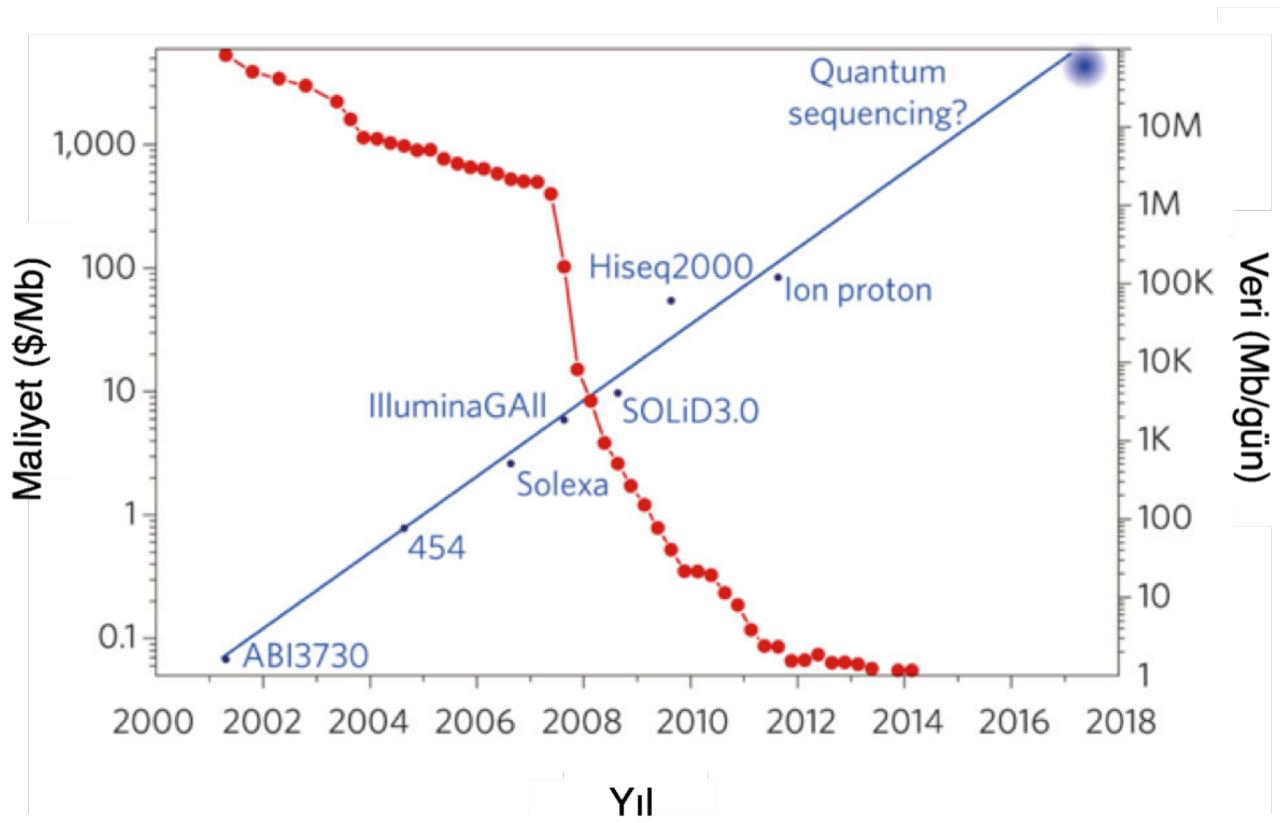
Sanger Sequencing



Genomik Devrim



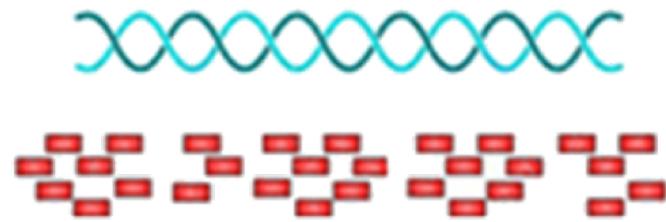
Genomik Devrim



Sanger vs NGS

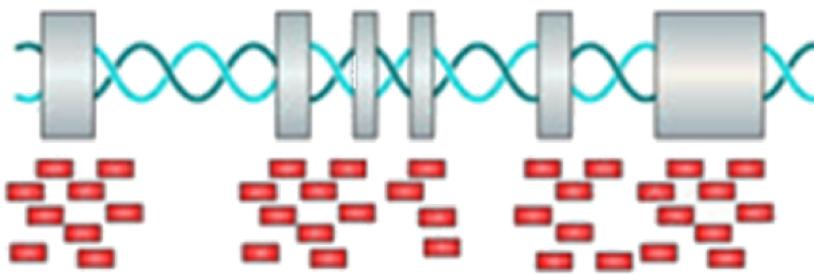
	Sanger	NGS
Örnek	PCR ürünü ya da klonlar	DNA kütüphaneleri
İşlem basamakları	Az sayıda, reax. sonrası temizleme	Çok ve karmaşık
Örnek sayısı	96 ya da 384	1- 16 +
Veri	1 örnek = 1 okuma	1 örnek = binlerce hatta milyonlarca okuma

Tüm Genom Dizileme (WGS)



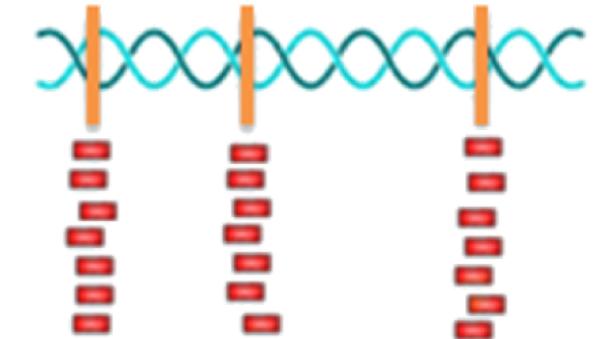
- Dizilenen bölge: tüm genom
- Derinlik $> 30X$
- Her tür variant tanımlanabilir (SNP, indel, SV vd)

Tüm Ekzom Dizileme (WES)



- Dizilenen bölge: tüm ekzom
- Derinlik $> 50 X \sim 100 X$
- Kodlayıcı bölgedeki her tür varyantı tanımlayabilir (SNP, indel, SV vd)
- Maliyeti düşük

Hedeflenmiş Dizileme

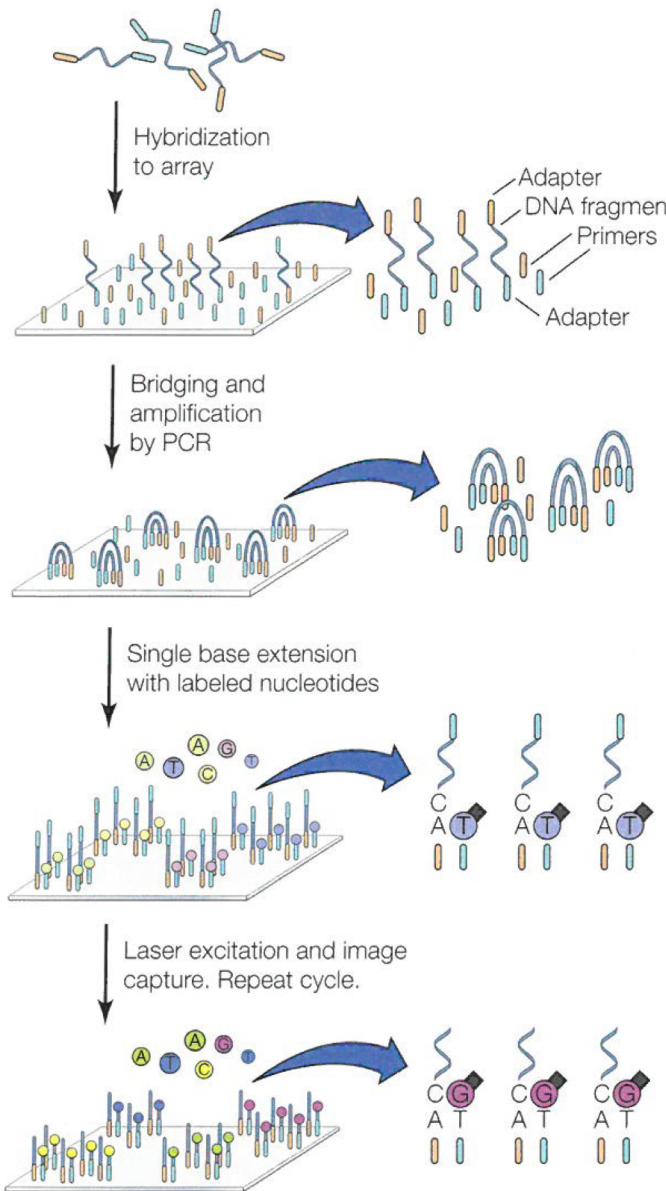


- Dizilenen bölge: özgül bölgeler
- Derinlik $> 500X$
- Ilgilenilen bölgedeki her tür varyantı tanımlayabilir (SNP, indel, SV vd)
- Maliyeti daha da düşük

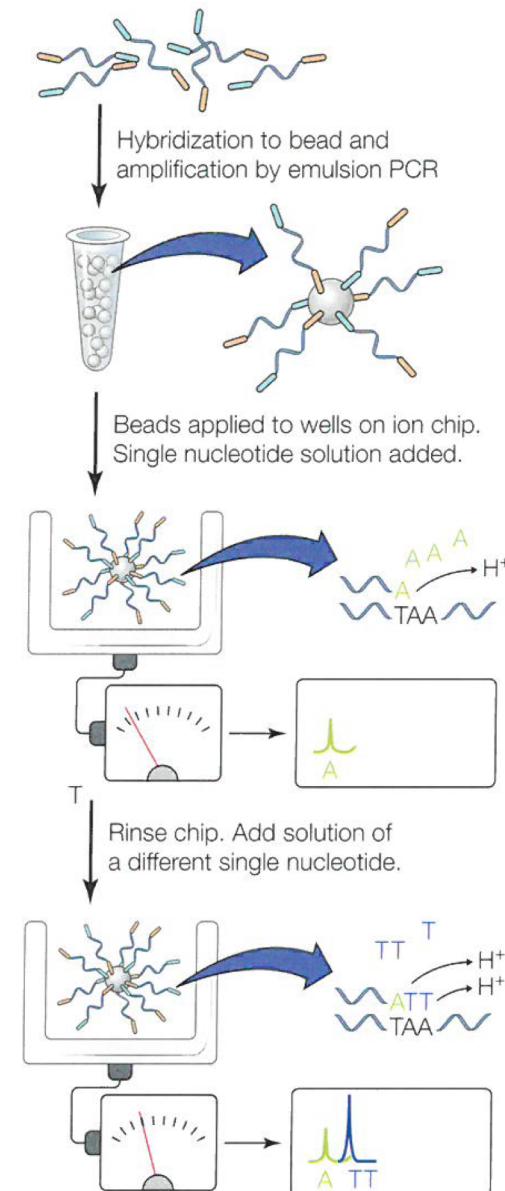
High-end sequencing- Platform[†]	Sequencing chemistry	Read lengths/through put	Run time	Template prep	Application
Roche 454 -Titanium FLX	Pyrosequencing	400 bp 400 Mb/run	10 hours	Emulsion PCR	Denovo WGS of microbes, pathogen discovery, Exome seq
Illumina/Solexa -HiSeq 2000	Reversible terminator chemistry	2×100bp 600 GB/run (dual cell)	11.5 days	Solid-phase	Human WGS, exome seq, RNA-seq, Methylation
ABI/LifeTechnology-SOLiD 5550XL	Sequencing by ligation	2×60bp 15 GB/day	8 days	Emulsion PCR	Human WGS, exome seq, RNA-seq, Methylation
HelicosBiotechnologies	Reversible Terminator chemistry	25-55 bp 28 GB/run (avg)	>1 GB/hour	Single molecule	Human WGS, exome seq, RNA-seq, Methylation
Roche 454- GS Junior	Pyrosequencing	400 bp 50 Mb/run	10 hours	Emulsion PCR	Denovo WGS of microbes, pathogen discovery, Exome seq
Illumina/Solexa- MiSeq	Reversible terminator chemistry	2×150bp 1.0-1.4 Gb	26 hours	Solid-phase	Microbial discovery, Exome seq, Targeted capture
ABI/Lifetechnology- Iontorrent	H+ Ion sensitive transistor	320 Mb/run	8 hours*	Emulsion PCR	Microbial discovery, Exome seq, Targeted capture

*Sample preparation – 6 hours, sequencing time – 2 hours, [†]Data shown here represent the highest figures currently available on the company website and is highly likely to change by the time this article is published

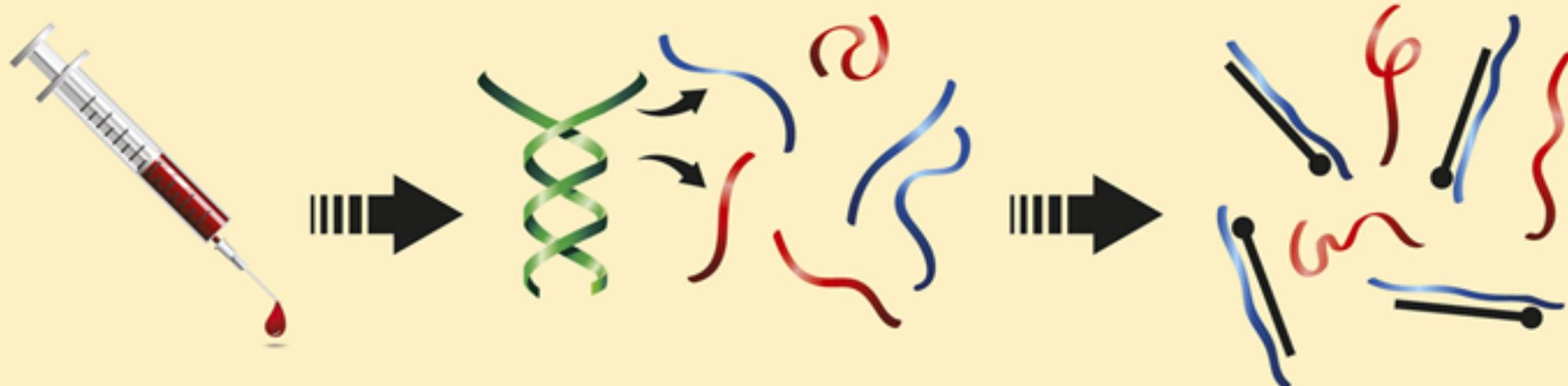
(A) Illumina's sequencing by synthesis



(B) Life Technologies' semiconductor sequencing



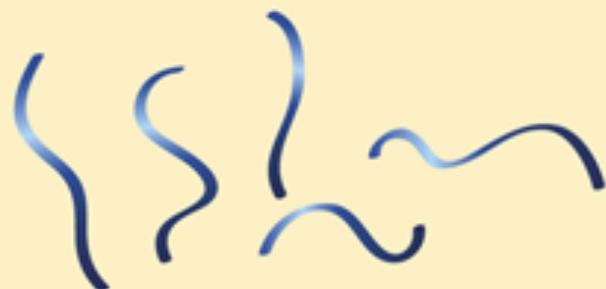
Tüm Ekzom Dizileme



1) Collect blood

2) Extract and fragment DNA

3) Capture exome DNA with probes



4) Recover only exome DNA fragments

5) Sequence on next-generation platform

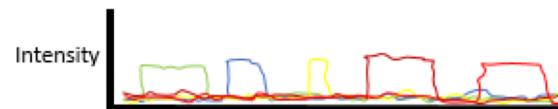
Yeni teknolojiler

PacBio
SMRT seq

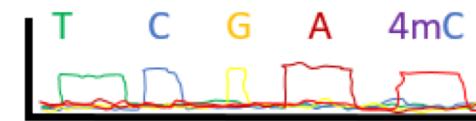
DNA passes thru
polymerase in an
illuminated volume



Raw output is fluorescent signal
of the nucleotide incorporation,
specific to each nucleotide

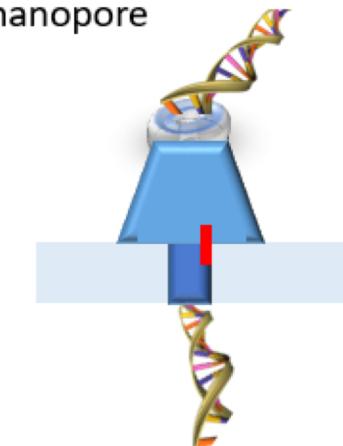


A,C,T,G have known pulse
durations, which are used to
infer methylated nucleotides

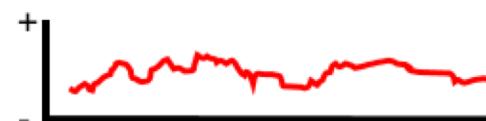


Oxford
Nanopore

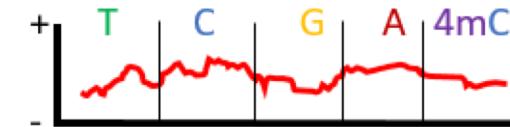
DNA passes thru
nanopore



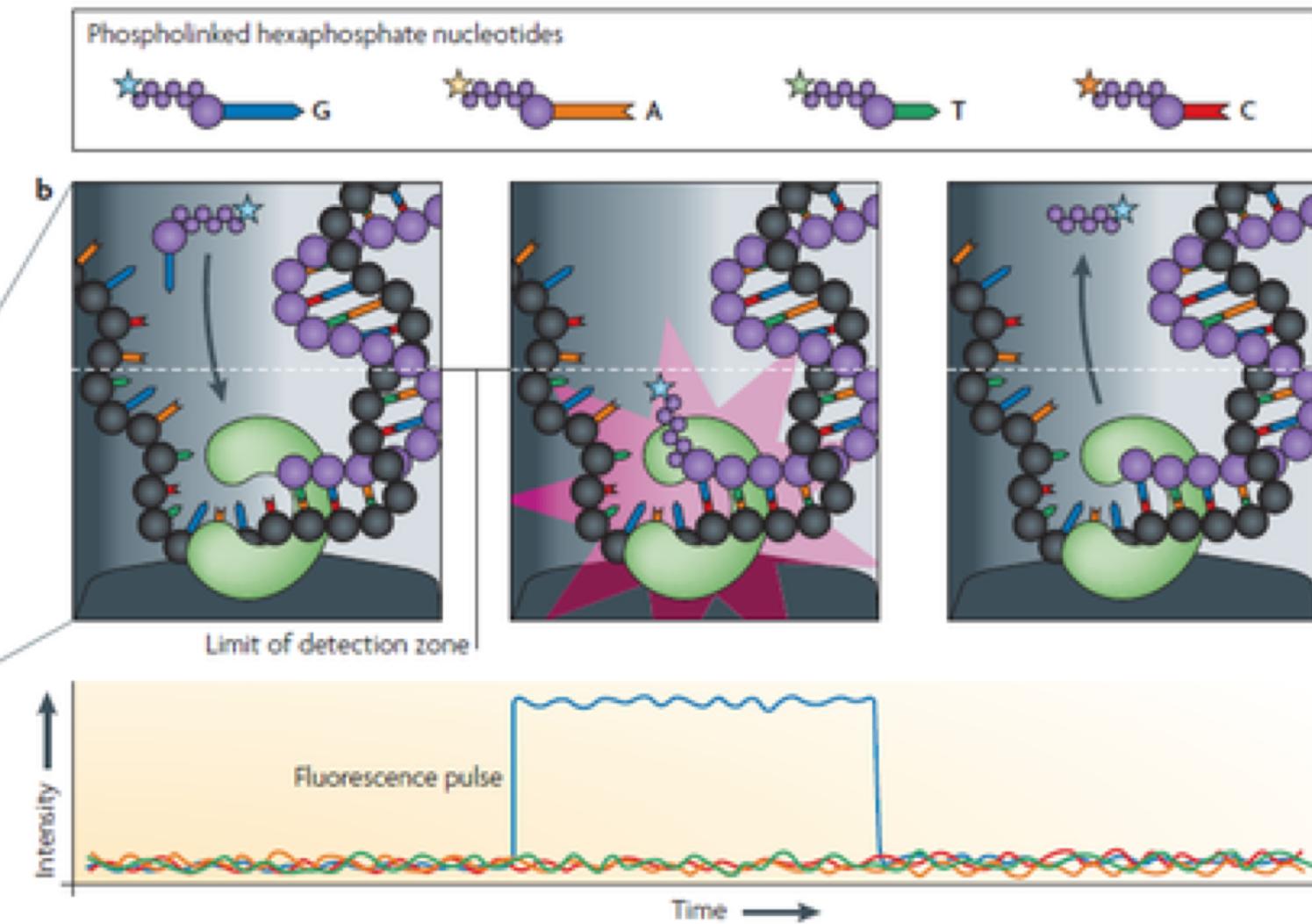
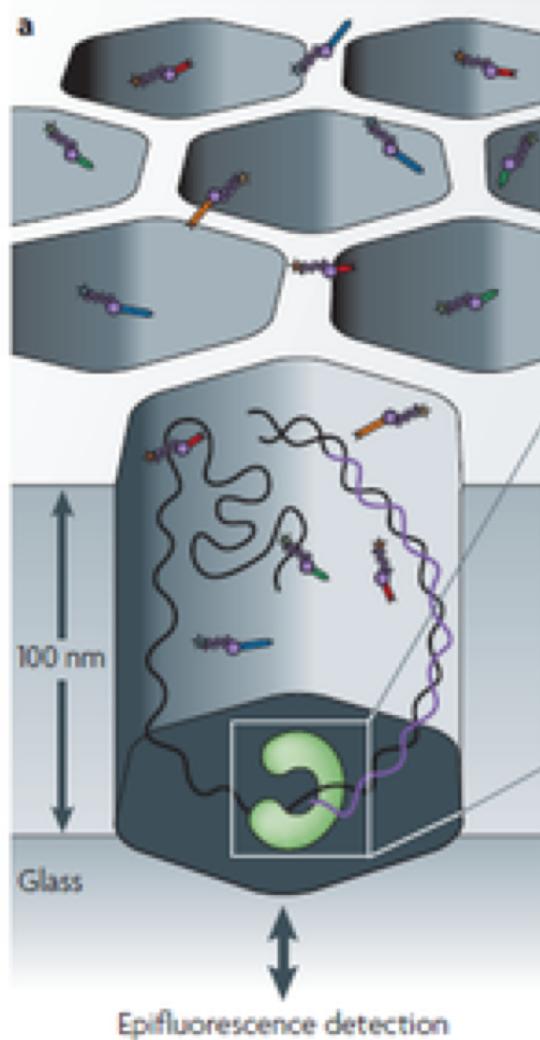
Raw output is electrical signal
caused by nucleotide blocking
ion flow in nanopore



Each nucleotide has a specific
electric "signature"

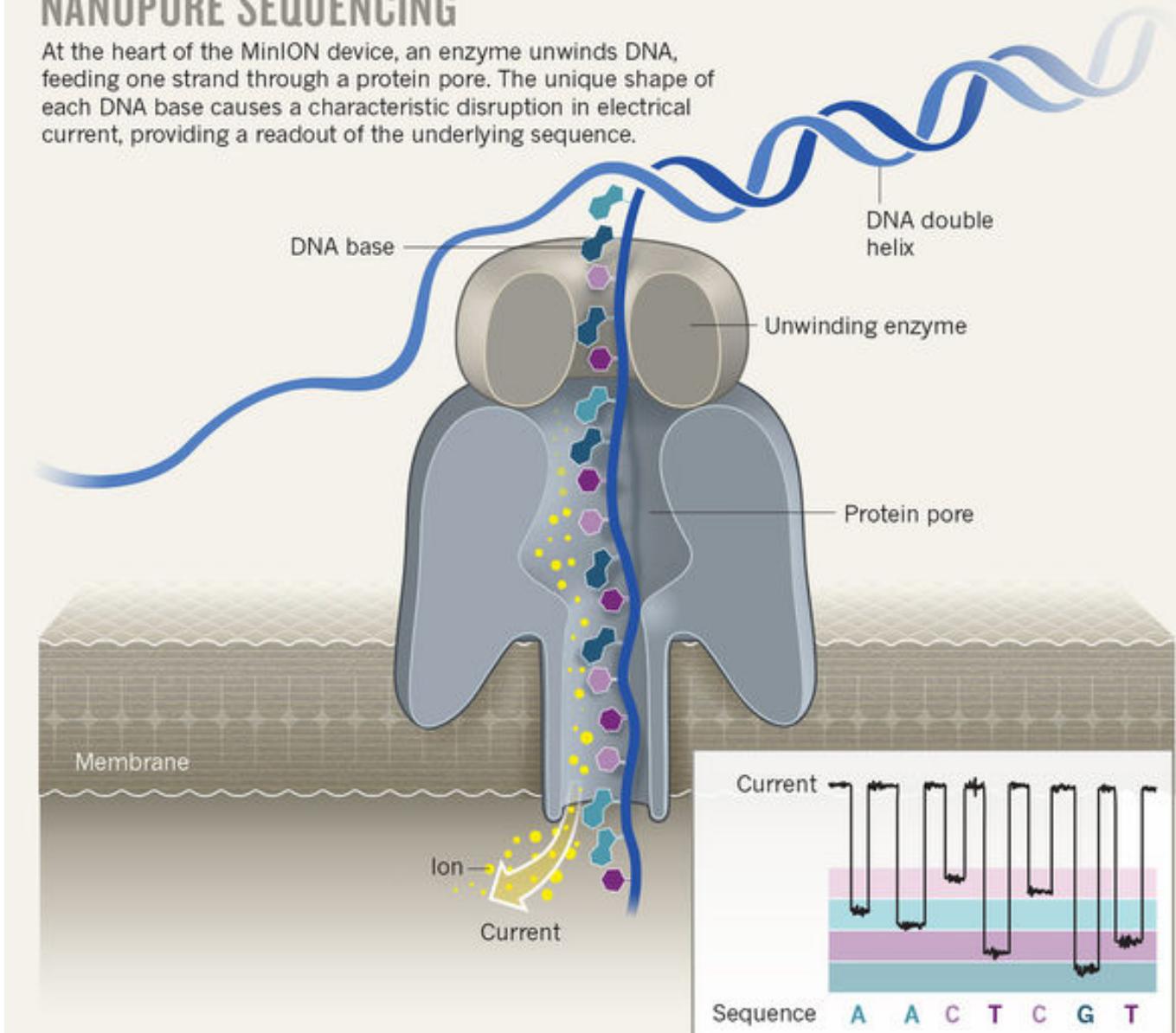


Pacific Biosciences — Real-time sequencing



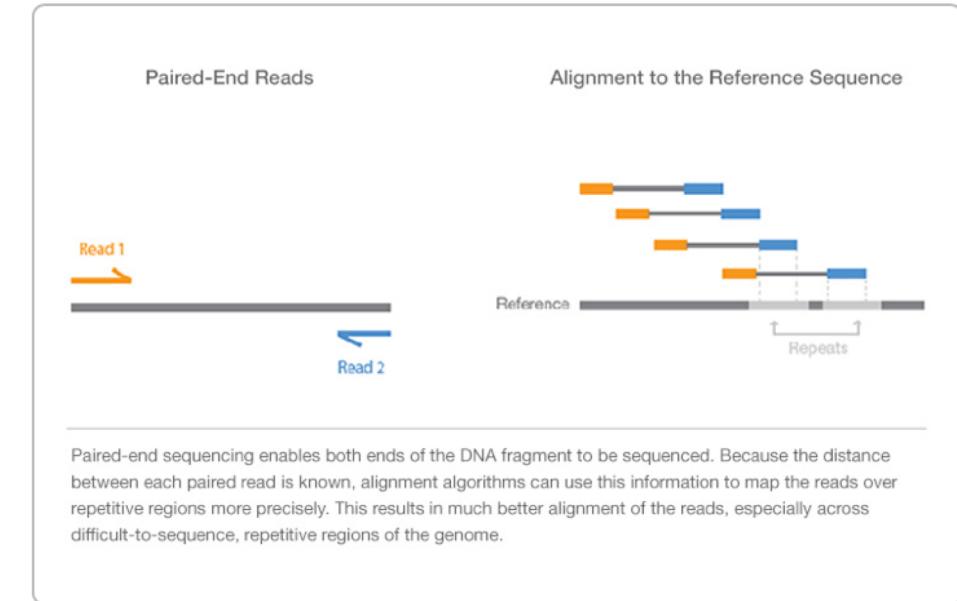
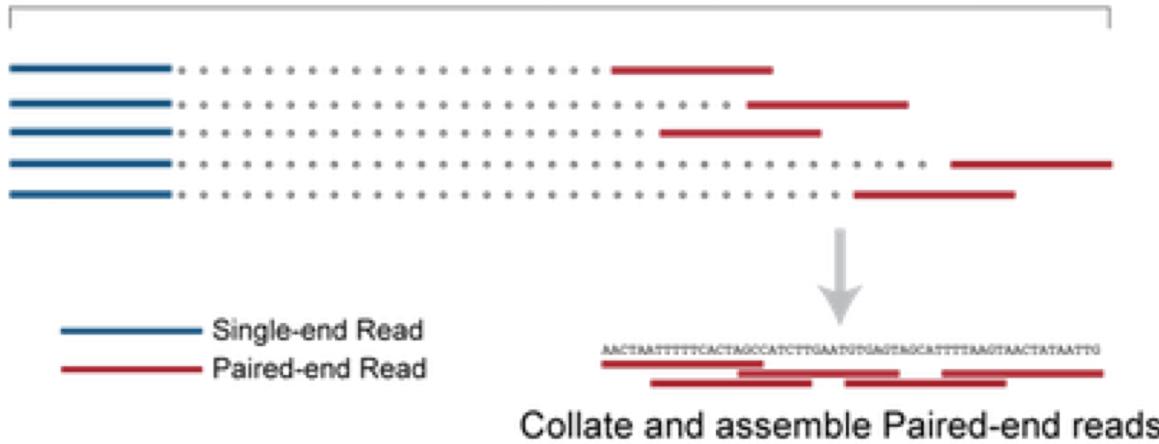
NANOPORE SEQUENCING

At the heart of the MinION device, an enzyme unwinds DNA, feeding one strand through a protein pore. The unique shape of each DNA base causes a characteristic disruption in electrical current, providing a readout of the underlying sequence.

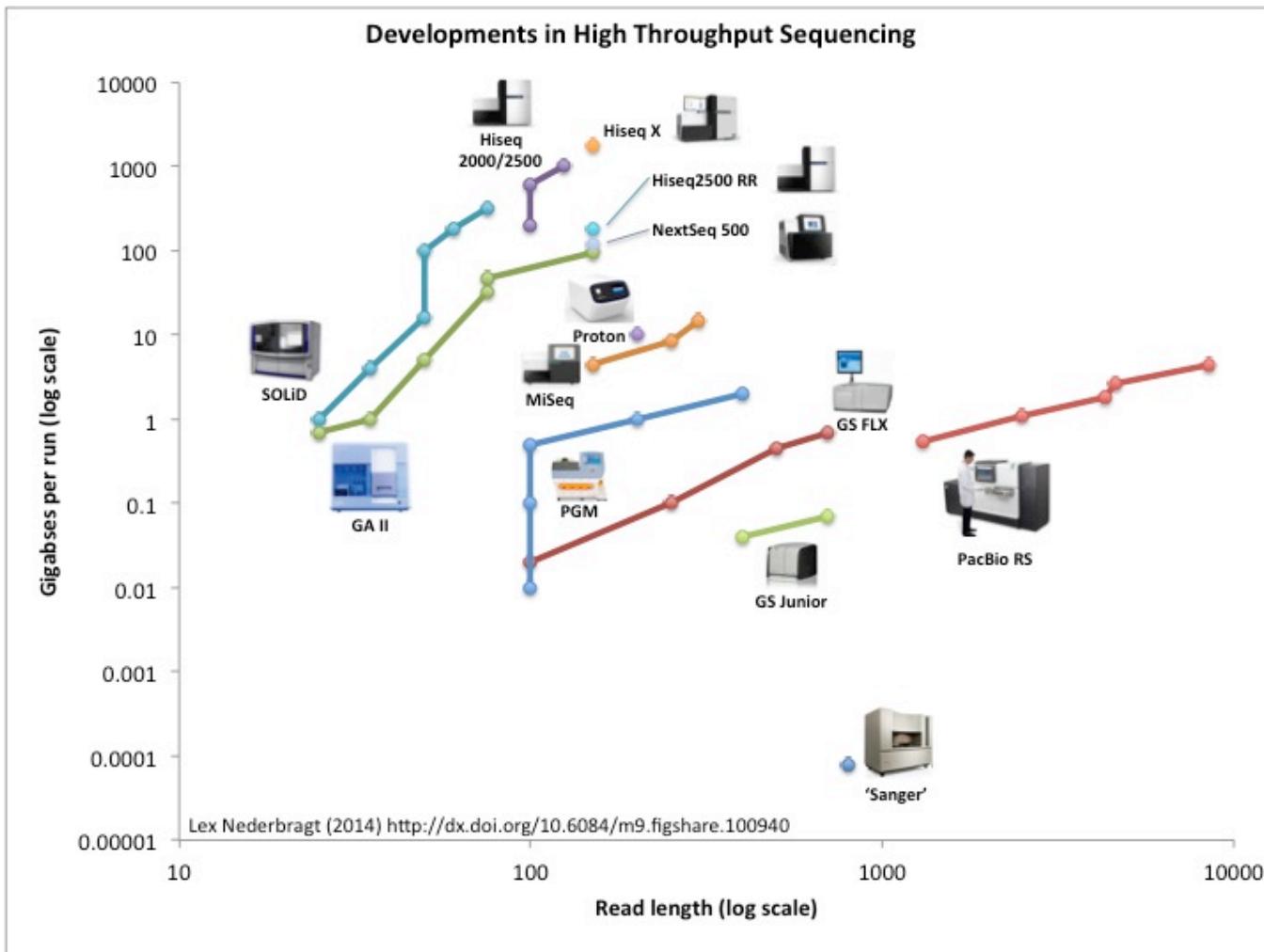


Kütüphane Özellikleri

Önemli noktalar: Tek uçtan okuma (single rad, SR) ya da çift yönlü okuma (paired end, PE)



Önemli noktalar: Okuma uzunluğu



Önemli noktalar: Okuma sayısı

Numbers of Single Reads by Instrument Manufacturer

Platform	Instrument	Unit	Reads / Unit	Reference
Illumina	HiSeq X Ten	Lane	375,000,000	1
Illumina	HiSeq 3000/4000	Lane	312,500,000	1
Illumina	HiSeq NextSeq 500 High-Output	Run	400,000,000	2
Illumina	HiSeq NextSeq 500 Mid-Output	Run	130,000,000	2
Illumina	HiSeq High-Output v4	Lane	250,000,000	3
Illumina	HiSeq High-Output v3	Lane	186,048,000	3
Illumina	HiSeq Rapid Run	Lane	150,696,000	3
Illumina	HiScanSQ	Lane	93,024,000	3
Illumina	GAIx	Lane	42,075,000	3
Illumina	MiSeq v3	Lane	25,000,000	4
Illumina	MiSeq v2	Lane	16,000,000	3
Illumina	MiSeq	Lane	5,000,000	3
Illumina	MiSeq v2 Micro	Lane	4,000,000	5
Illumina	MiSeq v2 Nano	Lane	1,000,000	5

Önemli noktalar: Kapsama (coverage)



Genom uzunluğu: L
Parçacık sayısı: n
Ortalama parçacık uzunluğu: l

Kapsama / dizileme derinliği (coverage / sequencing depth):
Genomdaki her baza denk gelen **ortalama** parçacık sayısı

Kapsama

$$K = n \cdot l / L$$

30X kapsama = baz başına ortalama 30 parçacık

Lander-Waterman modeli:

Tekbiçim dağılım (uniform distribution) varsayıımı ile, $K=10$ ise her 1 milyon nükleotidde 1 boş bölge kalır

Önemli noktalar: Kapsama (coverage)

Estimate of Coverage Requirements by Application Type

Application Type	Coverage
DNA-Seq (Re-Sequencing)	30 - 80X
DNA-Seq (De novo assembly)	100X
SNP Analysis / Rearrangement Detection	10 - 30X
Exome	100 - 200X
ChIP-Seq	10 - 40X

Farklı uygulamalar için kapsama ve okuma önerileri

Category	Detection or Application	Recommended Coverage (x) or Reads (millions)	References
Whole genome sequencing	Homozygous SNVs	15x	Bentley et al., 2008
	Heterozygous SNVs	33x	Bentley et al., 2008
	INDELs	60x	Feng et al., 2014
	Genotype calls	35x	Ajay et al., 2011
	CNV	1-8x	Xie et al., 2009; Medvedev et al., 2010
Whole exome sequencing	Homozygous SNVs	100x (3x local depth)	Clark et al., 2011; Meynert et al., 2013
	Heterozygous SNVs	100x (13x local depth)	Clark et al., 2011; Meynert et al., 2013
	INDELs	not recommended	Feng et al., 2014
Transcriptome Sequencing	Differential expression profiling	10-25M	Liu Y. et al., 2014; ENCODE 2011 RNA-Seq
	Alternative splicing	50-100M	Liu Y. et al., 2013; ENCODE 2011 RNA-Seq
	Allele specific expression	50-100M	Liu Y. et al., 2013; ENCODE 2011 RNA-Seq
	De novo assembly	>100M	Liu Y. et al., 2013; ENCODE 2011 RNA-Seq

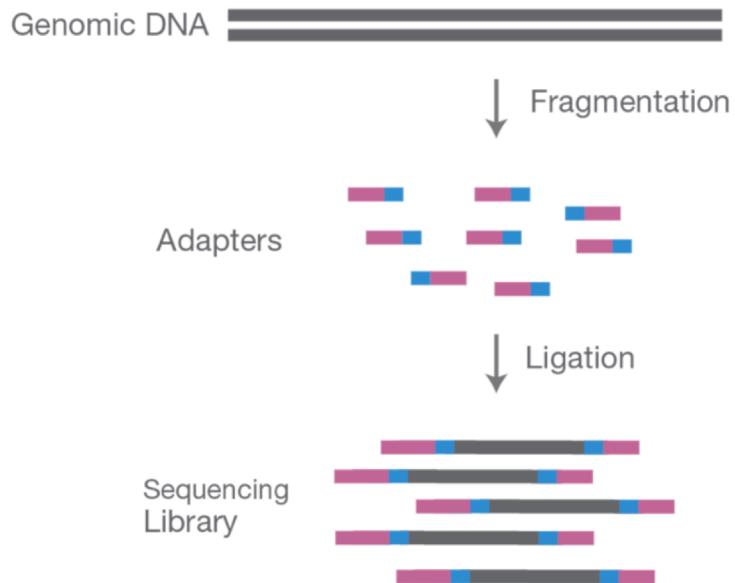
Yeni nesil DNA dizileme basamakları

- Amaca uygun yöntemin seçilmesi
- Kütüphane hazırlama
- Küme oluşturma
- Dizileme
- Veri analizi

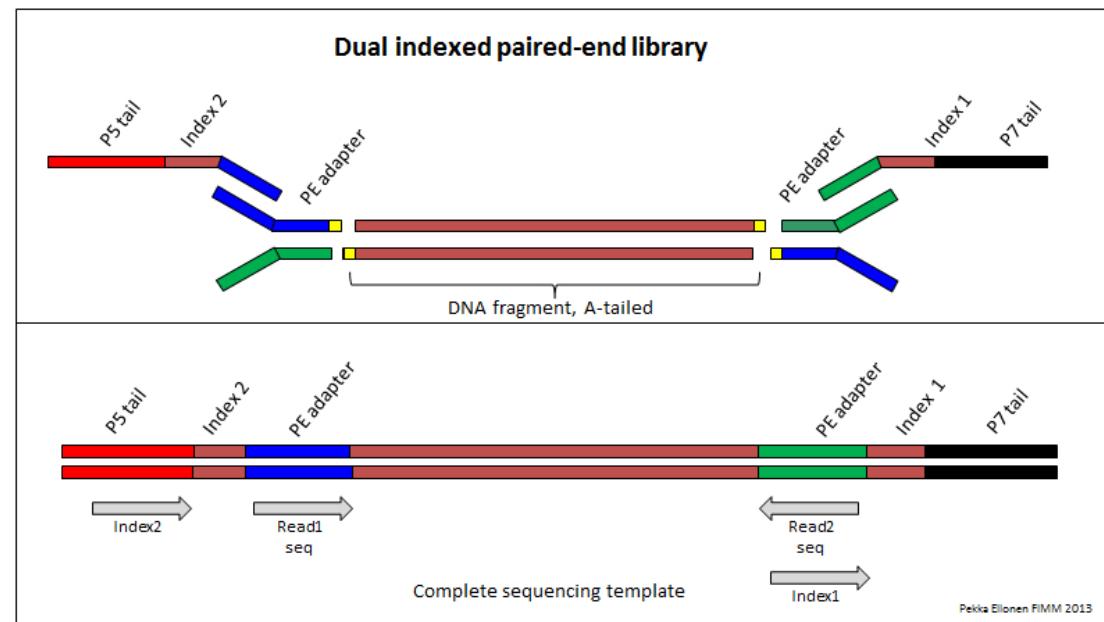
Kütüphane hazırlama: DNA fragmentasyonu

- Fiziksel (akustik, sonikasyon, hidrodinamik)
- Kimyasal (sıcaklık, divalent metal)
- Enzimatik (Dnase I ya da diğer özgül olmayan nükleazlar, transpozaz)

A. Library Preparation

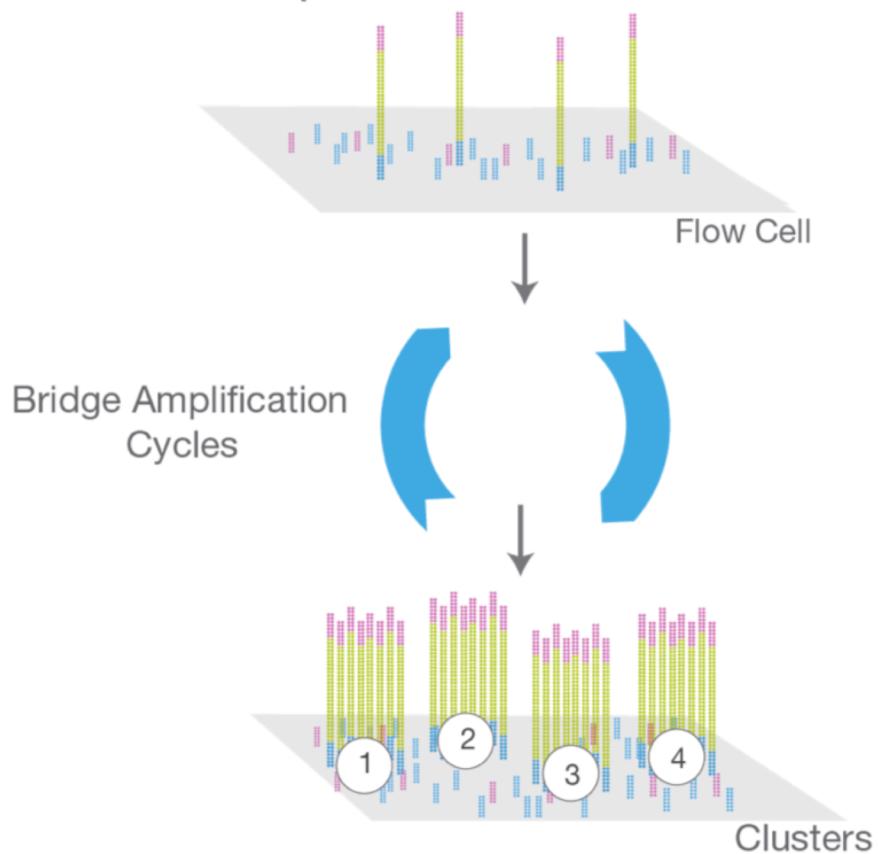


NGS library is prepared by fragmenting a gDNA sample and ligating specialized adapters to both fragment ends.



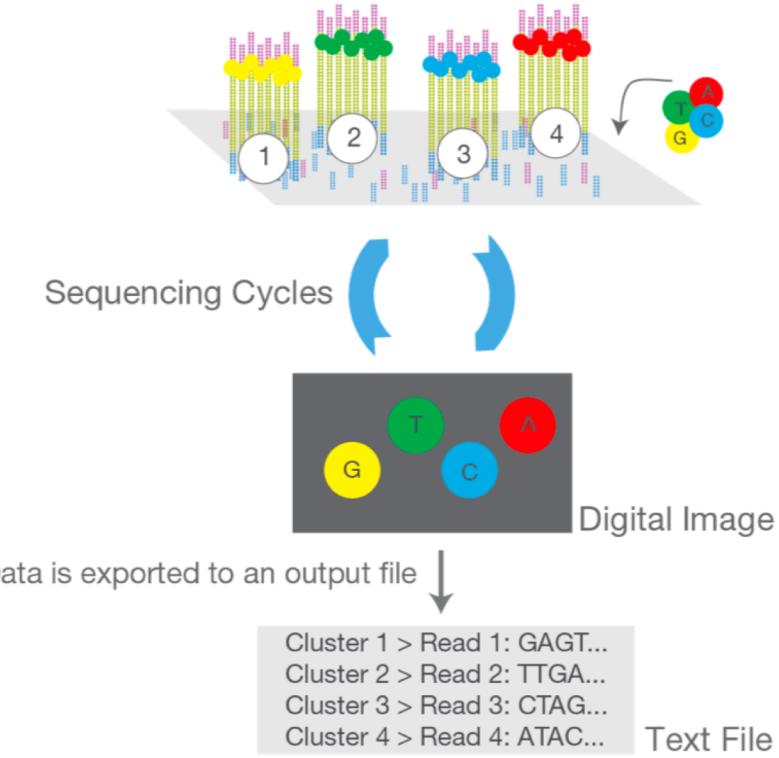
Pekka Ilonen FIMM 2013

B. Cluster Amplification



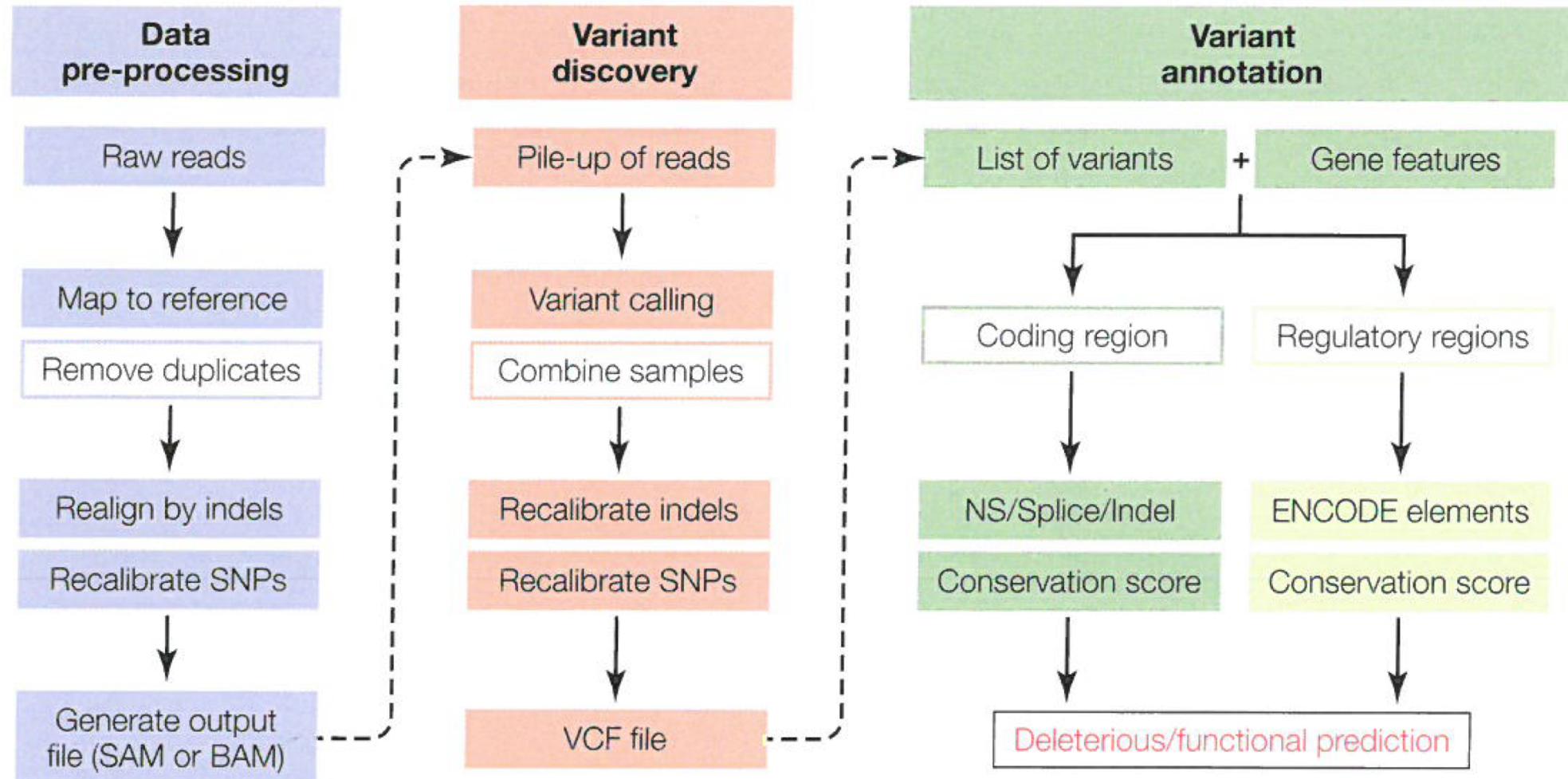
Library is loaded into a flow cell and the fragments are hybridized to the flow cell surface. Each bound fragment is amplified into a clonal cluster through bridge amplification.

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.

Verilerin analizi



Bazı durumlar deneyinizi kötü etkileyebilir:

- Haritalanmayan okumalar
- PCR duplikasyonları
- Düşük kaliteli okumalar
- Adaptör dimeri ya da bu dimerlerin dizilenmesi
- Özgün haritalanmayan okumalar
- Bilgi vermeyen dizilere haritalanan okumalar (rRNA genleri gibi)

Son olarak, NGS özellikleri

- Kısa okumalar
 - ~500 bp (Roche 454)
 - 35-150 bp (Illumina, AB)
- Çok fazla miktarda dizi ve okuma eldesi
 - 600 Gbp/run (Illumina)
- Yüksek ve düşük GC içeriği etkisi
- Yüksek hata oranı