

Diagnoseq Doku Tipleme Laboratuvarı

M.Sc. Zekiye Cemre Demir

Biyoinformatik Laboratuvar Yöneticisi



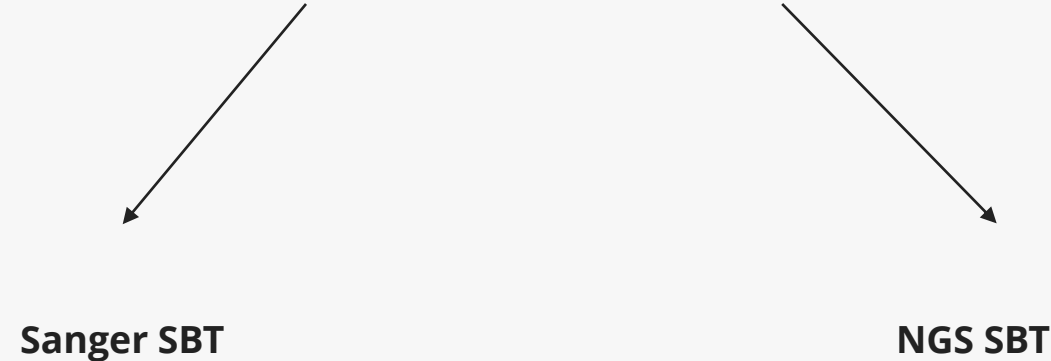
Transplantasyonda NGS Uygulamaları

- Yüksek Çözünürlüklü Tiplendirme
- Kimerizm
- cfDNA NGS uygulaması
- NGS KIR Tiplendirme
- NGS MICA, MICB Tiplendirme

Yeni Nesil Dizileme Genel Bakış

SBT Terminolojisi

Sequencing Based Typing



```
graph TD; A[Sequencing Based Typing] --> B[Sanger SBT]; A --> C[NGS SBT]
```

Sanger SBT

NGS SBT

Yeni Nesil Dizileme Genel Bakış

1. Nesil Dizileme

- Birden çok benzer moleküllerin dizilenmesi
- Büyük jellerde ya da kapiller sistemlerde dizilemenin limiti

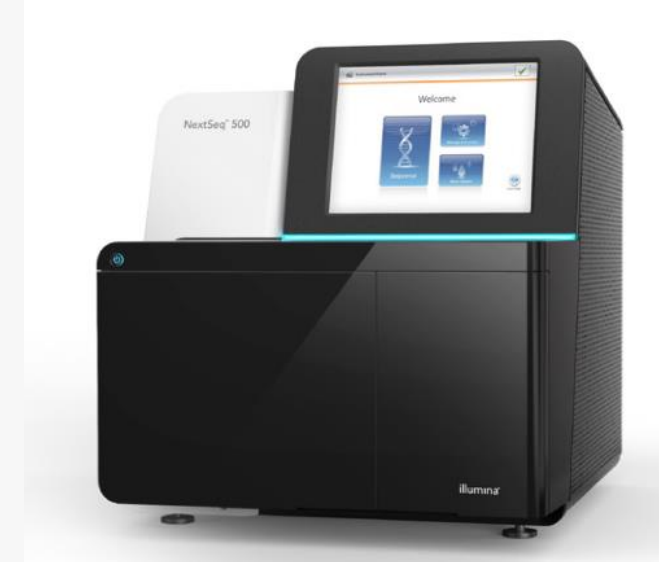
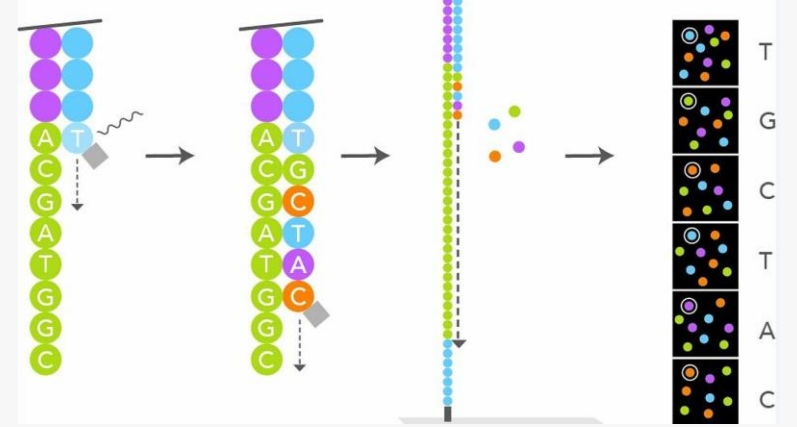
2. Nesil Dizileme

- Milyonlarca klonal amplikonların dizilenmesi
- Bir yüzey alanında fragmentlerin sabitlenmesi



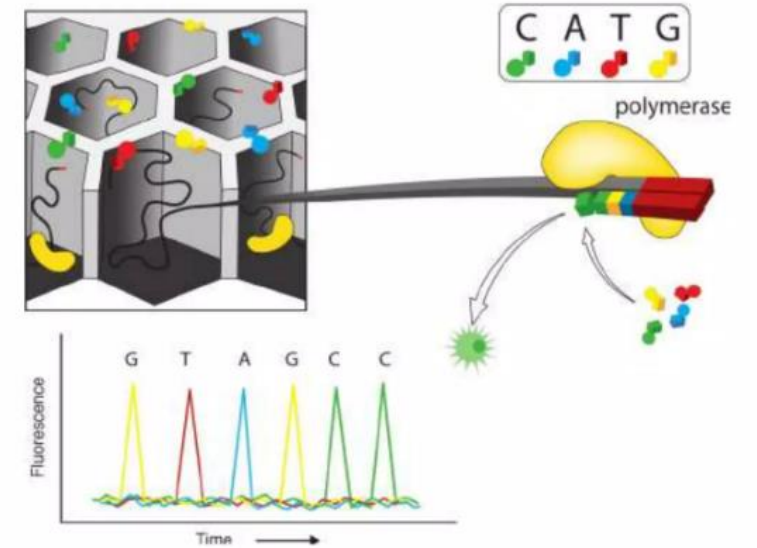
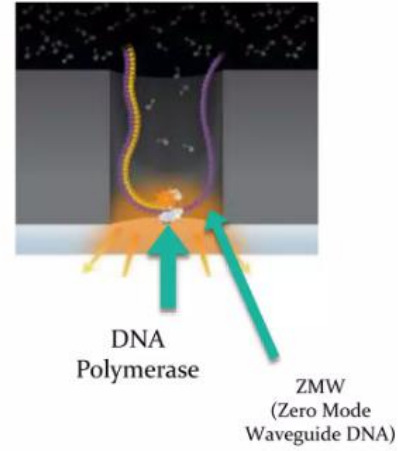
Yeni Nesil Dizileme Genel Bakış

- Yüksek hacimli DNA Dizileme
- Zaman ve paradan kazanç
- Binlerce diziye tek seferde dizileme
- Yüksek ölçüde data üretilmesi



3. Nesil Dizileme Teknolojileri

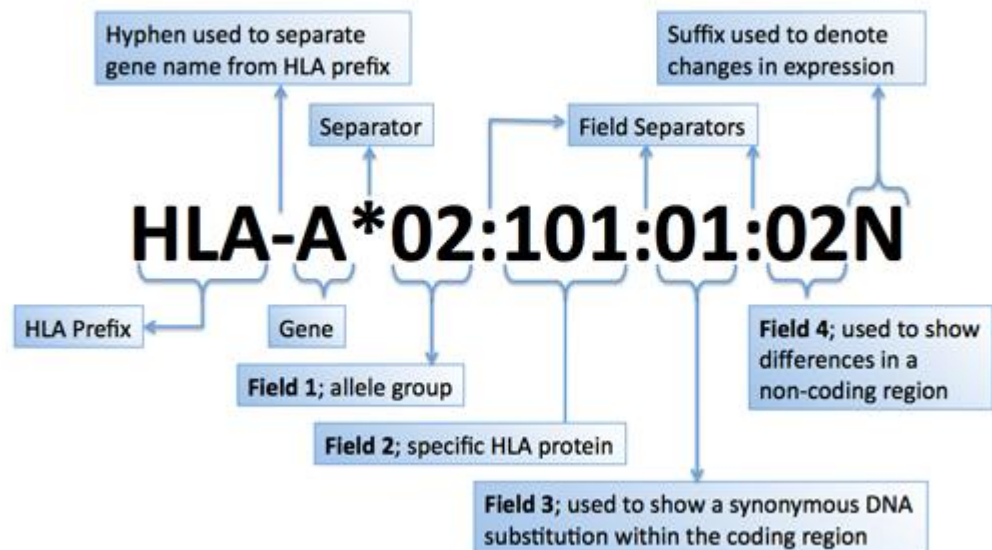
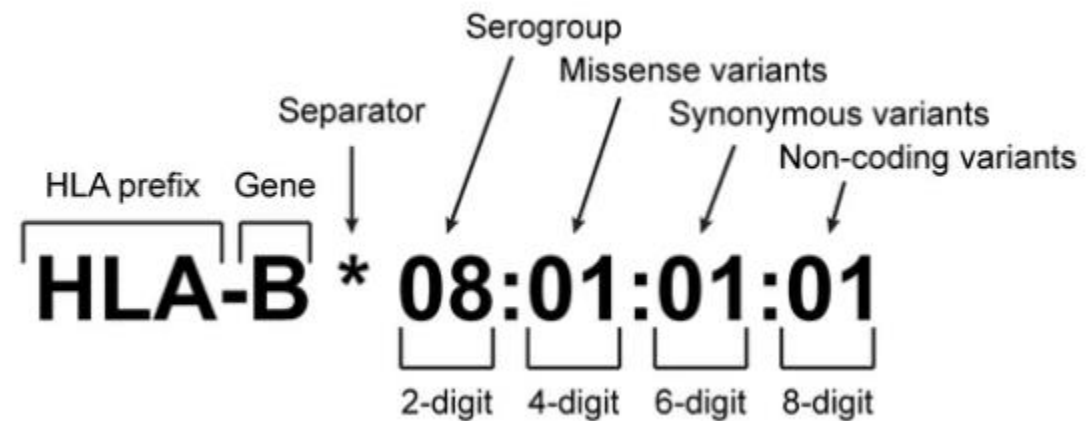
- Tek molekül dizileme
- DNA amplifikasyonu yok
- Uzun okuma
- Düşük coverage



- Roche/454 > 400bp
- Illumina/Solexa > 100bp
- Pacific Bioscience > 1000 bp and single molecule

HLA Nomenclature

- HLA genlerindeki polimorfizmin yüksek olması
- Her allelin özgün bir numarası vardır.
- Totalde 4 set digite kadar isimlendirme alabilir (8 digit, 4 alan)



G Kodu (G Group)

Class 1 HLA genleri peptid kodlama bölgesi (peptide binding-domain) ekzon 2 ve ekzon 3

Class 2 HLA genleri peptid kodlama (peptide binding-domain) bölgesi ekzon 2

G kodu ya da G grubu peptid kodlama bölgesinde aynı nükleotid dizisine sahip olan HLA allelleri 3. alandan itibaren G harfini alarak bir grubu temsil ederler.

P Kodu (P Group)

Class 1 HLA genleri protein kodlama bölgesi ekzon 2 ve ekzon 3

Class 2 HLA genleri protein kodlama bölgesi ekzon 2

P kodu ya da P grubu antijen bağlanma bölgelerindeki aynı protein dizisine sahip allelleri temsil eder.

<https://hla.alleles.org/nomenclature/index.html>

<https://www.imgt.org/>

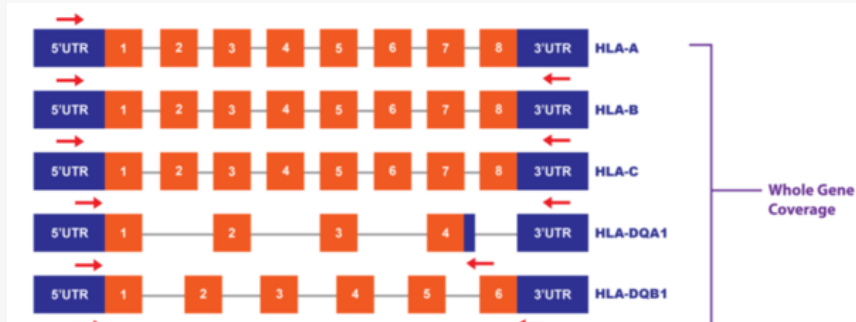
<https://www.ebi.ac.uk/ipd/imgt/hla/>

Yeni Nesil Dizileme İş Akışı – Hedefe Yönelik

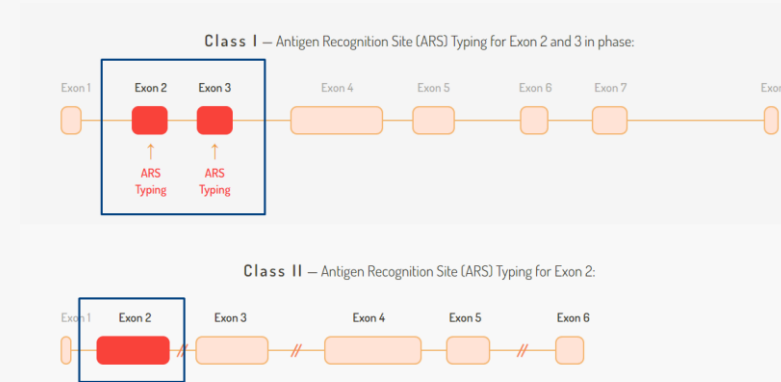
Hedeflenen genlerin çoğaltılması (PCR)



Long- range PCR metodu (tüm gen çoğaltılması)



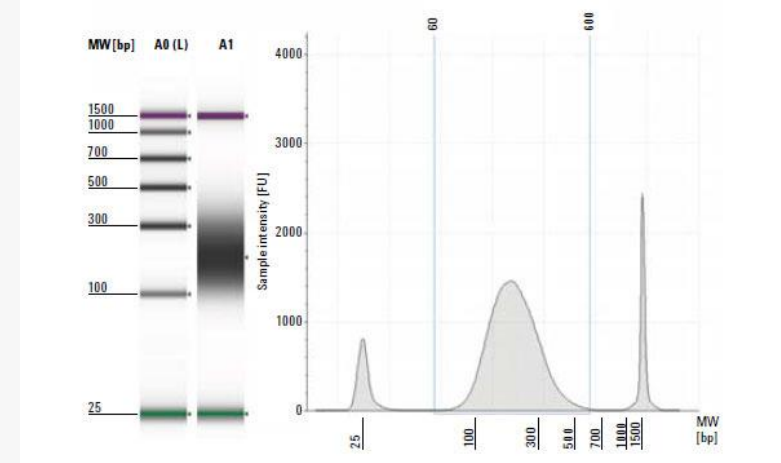
Amplikon tabanlı çoğaltma (gendeki belirli bölgelerin çoğaltılması)



Yeni Nesil Dizileme İş Akışı – Hedefe Yönelik

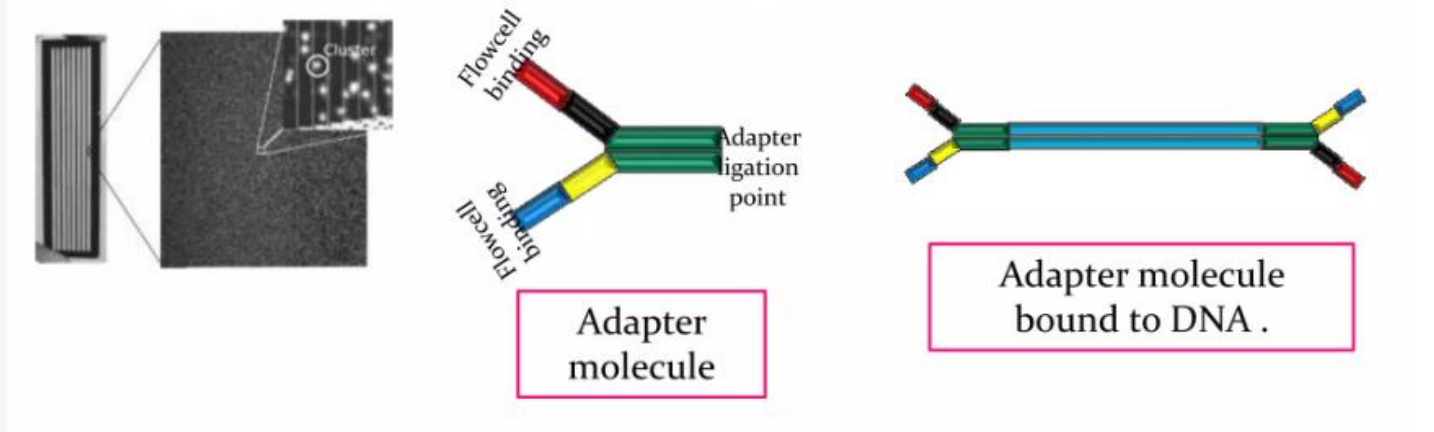
Büyükölge göre fragment seçilimi (*Size-Selection*)

- Jel sistemleri ile seçim
- Bioanalyzer veya jel sistemleri büyüklük kontrolü



Yeni Nesil Dizileme İş Akışı – Hedefe Yönelik

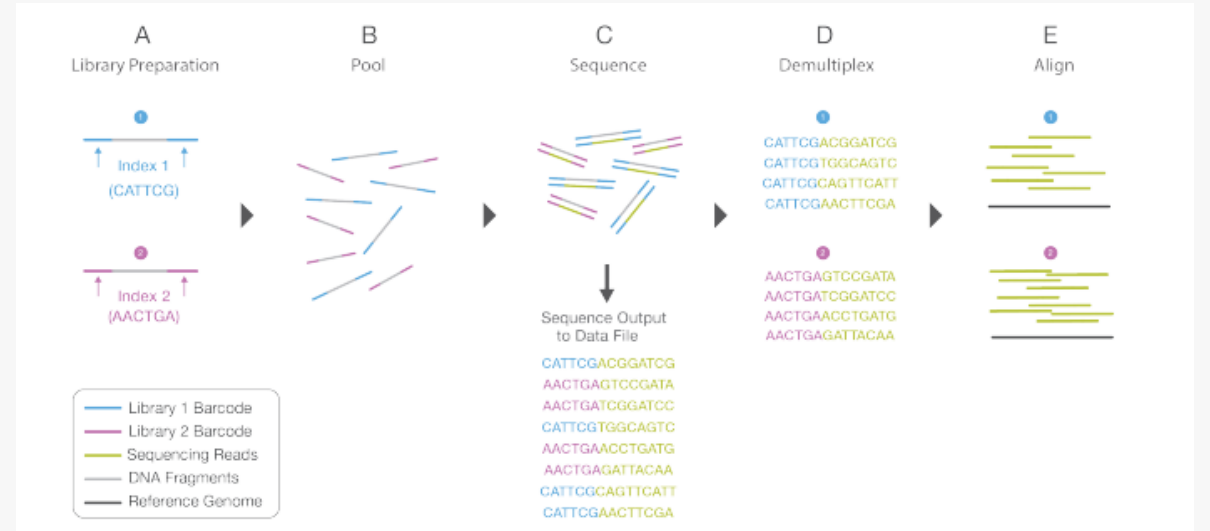
Platform spesifik adaptörlerin eklenmesi



- Flowcell bağlanmasını sağlayan adaptör
- Örnek ayırımını sağlayan indeksleme adaptörleri (örnek barkodlama)

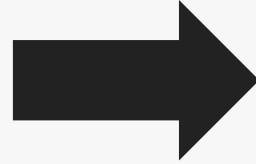
Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- Referansa hizalama (*mapping –alignment*)
- De novo assembly (referansa bağlı kalmadan hizalama)
- Phasing

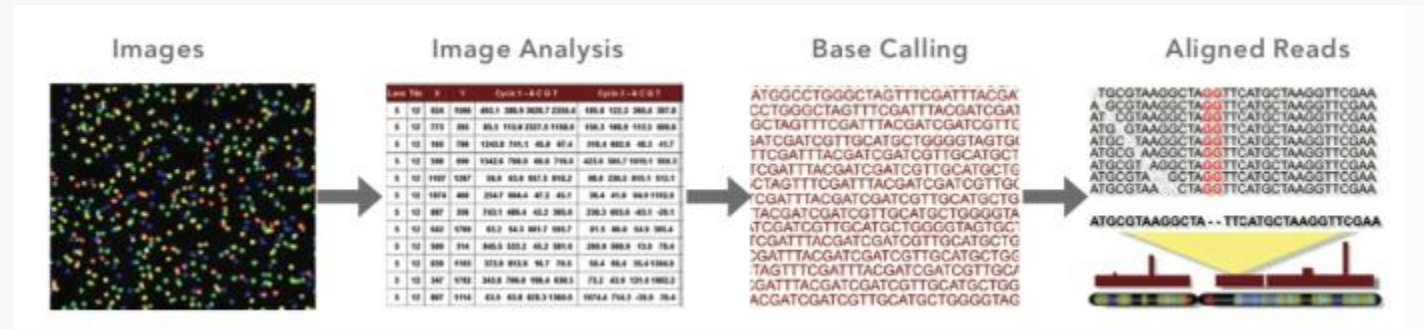


Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- Referansa hizalama (*mapping –alignment*)
- De novo assembly (referansa bağlı kalmadan hizalama)
- Phasing



Oluşan her clusterdan gelen sinyallerinden görüntü analizi ile nükleotidler atanır



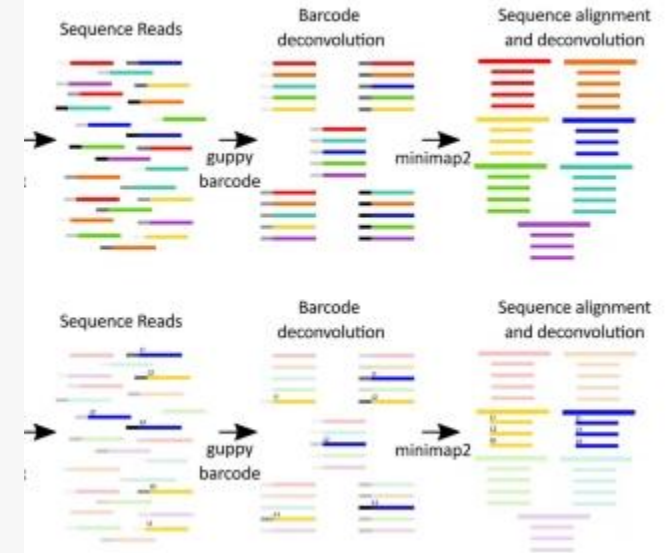
Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- Referansa hizalama (*mapping –alignment*)
- De novo assembly (referansa bağlı kalmadan hizalama)
- Phasing

Her biri farklı barkodlanmış örnekler fastq çevirimi sırasında okumaların örneklere atanmasını sağlar.

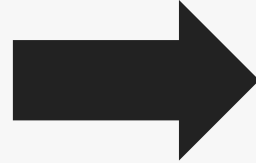


[Header]									
IEMFileVer		5							
Date		10.04.2023							
Workflow		GenerateFASTQ							
Application		NovaSeq FASTQ Only							
Instrument		NovaSeq							
Assay		QIAseq FX DNA							
Index Adap		QIAseq FX DNA UDI (96 Indexes) Set A							
Chemistry		Amplicon							
[Reads]									
149									
149									
[Settings]									
Adapter		AGATCGGAAGAGCACACGTCTGAACTCCAGTCA							
AdapterRe		AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT							
[Data]									
Sample_ID	Sample_N	Sample_P	Sample_V	Index_Plati	I7_Index	I1_Index	I5_Index	I10_Index2	Sample_P>Description
W00701213	A01	A01	UD1001	TGAACGT	UD1001	AGTCGGCCAT			
W00701214	B01	B01	UD1002	ACCAGAC	UD1002	GTGCTCATCG			
W00701215	C01	C01	UD1003	ACTGGCC	UD1003	TCGACTTATC			
W00701216	D01	D01	UD1004	GCCTTAC	UD1004	CAAGGCCGTA			
W00701217	E01	E01	UD1005	TTATCGG	UD1005	TTGTGTCCT			
W00701218	F01	F01	UD1006	GAGTAT	UD1006	GCCTACTGAG			
W00701219	G01	G01	UD1007	TCAAGGA	UD1007	CAGGCACCTC			
W00701220	H01	H01	UD1008	CGAACCC	UD1008	AGGCGAGAGA			
W00701221	A02	A02	UD1009	GAGCCAA	UD1009	GAAGGTGCCT			
W00701222	B02	B02	UD1010	AAGGCC	UD1010	TTACCAACAG			
W00701223	C02	C02	UD1011	TTAGAGA	UD1011	AGGTACCAGC			
W00701224	D02	D02	UD1012	TCTAAGA	UD1012	CCGCTCCTTA			
W00701225	E02	E02	UD1013	TGTAACC	UD1013	TGGAGCGATT			
W00701226	F02	F02	UD1014	CCGACAC	UD1014	CAATTAGGAG			
W00701227	G02	G02	UD1015	CTCTGAT	UD1015	ATTATGAGGC			
W00701228	H02	H02	UD1016	CGGCCT	UD1016	GCTCAATACA			
W00701229	A03	A03	UD1017	TGCATAG	UD1017	TGTTATGGCT			
W00701230	B03	B03	UD1018	AACTTC	UD1018	CCACTTGTGG			
W00800703	C03	C03	UD1019	AAGAGAT	UD1019	GTGTGATAC			
W00800713	D03	D03	UD1020	GCCTGAA	UD1020	AAGAACGGTA			



Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- **Referansa hizalama (*mapping -alignment*)**
- De novo assembly (referansa bağlı kalmadan hizalama)
- Phasing

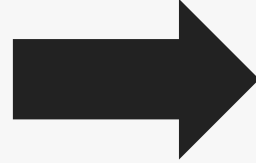


Dizilenen örneklerin kaynağına göre kullanılacak referans genoma üretilen fragmentler algoritmalar yardımı ile hizalanır.



Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- Referansa hizalama (*mapping –alignment*)
- De novo assembly (referansa bağlı kalmadan hizalama)
- Phasing



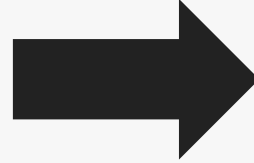
Popüler aligner'lar BOWTIE ve BWA

BAM dosyası

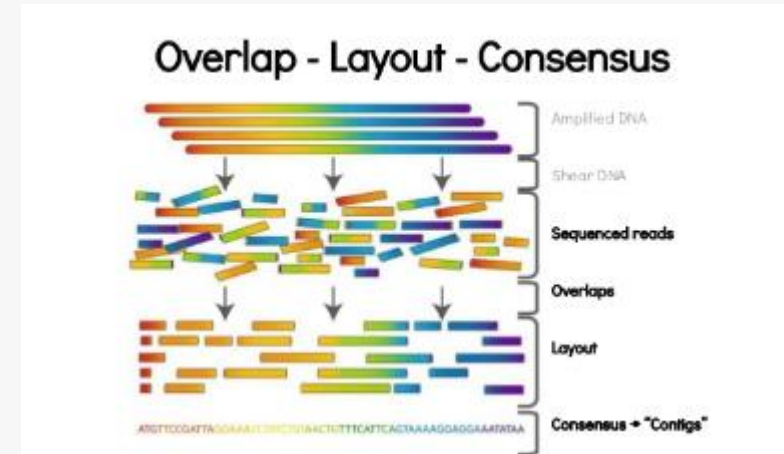


Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- Referansa hizalama (*mapping –alignment*)
- **De novo assembly (referansa bağlı kalmadan hizalama)**
- Phasing



Üretilen fragmentler herhangi bir referansa hizalamadan, overlap noktalarına göre hizalanarak bir konsensus dizi elde edilir.

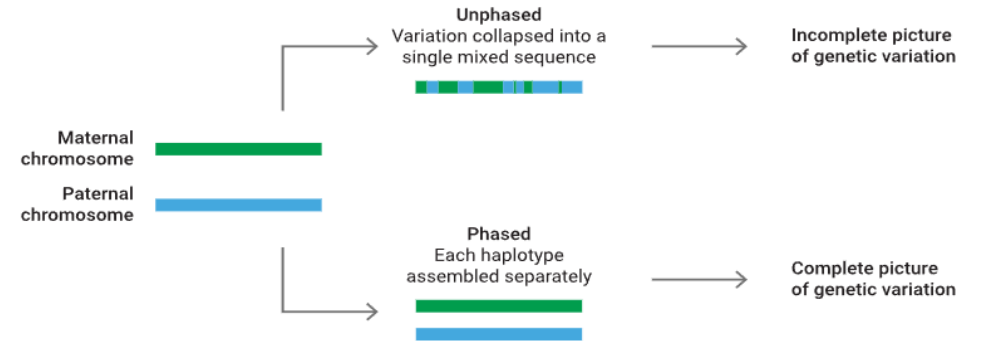
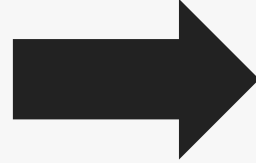


Yeni Nesil Dizileme İş Akışı – Veri Analizi

- Base calling
- Demultiplexing
- Referansa hizalama (*mapping -alignment*)
- De novo assembly (referansa bağlı kalmadan hizalama)
- **Phasing**

-----A--C-----A-----T--T-----
-----A--T-----A-----C--G-----

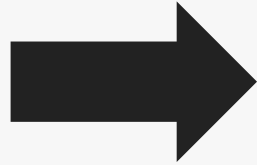
Anne babadan gelen her bir kopyanın karışmış şekilde gelen veri içerisinde istatistiksel modeller ile haplotip geçişlerinin sağlanması



Yeni Nesil Dizileme İş Akışı – Veri Analizi

Phasing Metodları

- 
- Base calling
 - Demultiplexing
 - Referansa hizalama (*mapping –alignment*)
 - De novo assembly (referansa bağlı kalmadan hizalama)
 - **Phasing**



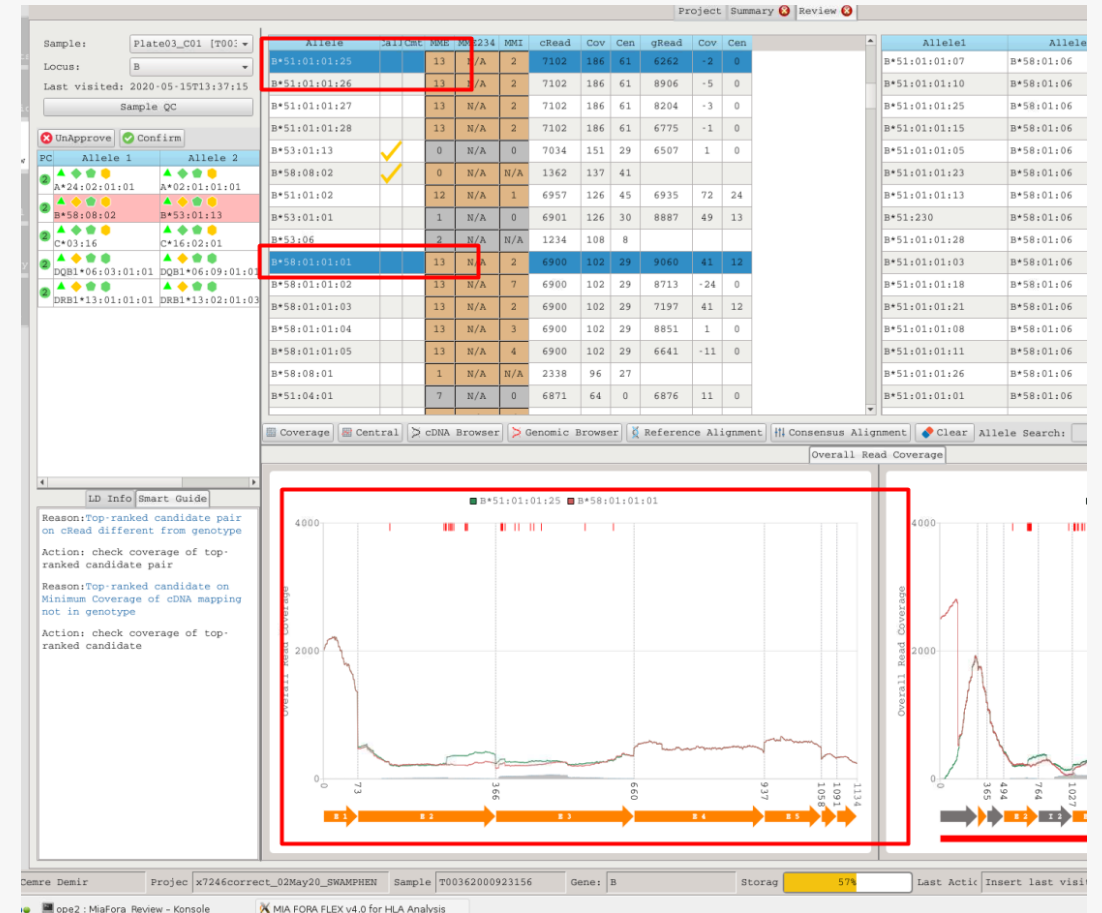
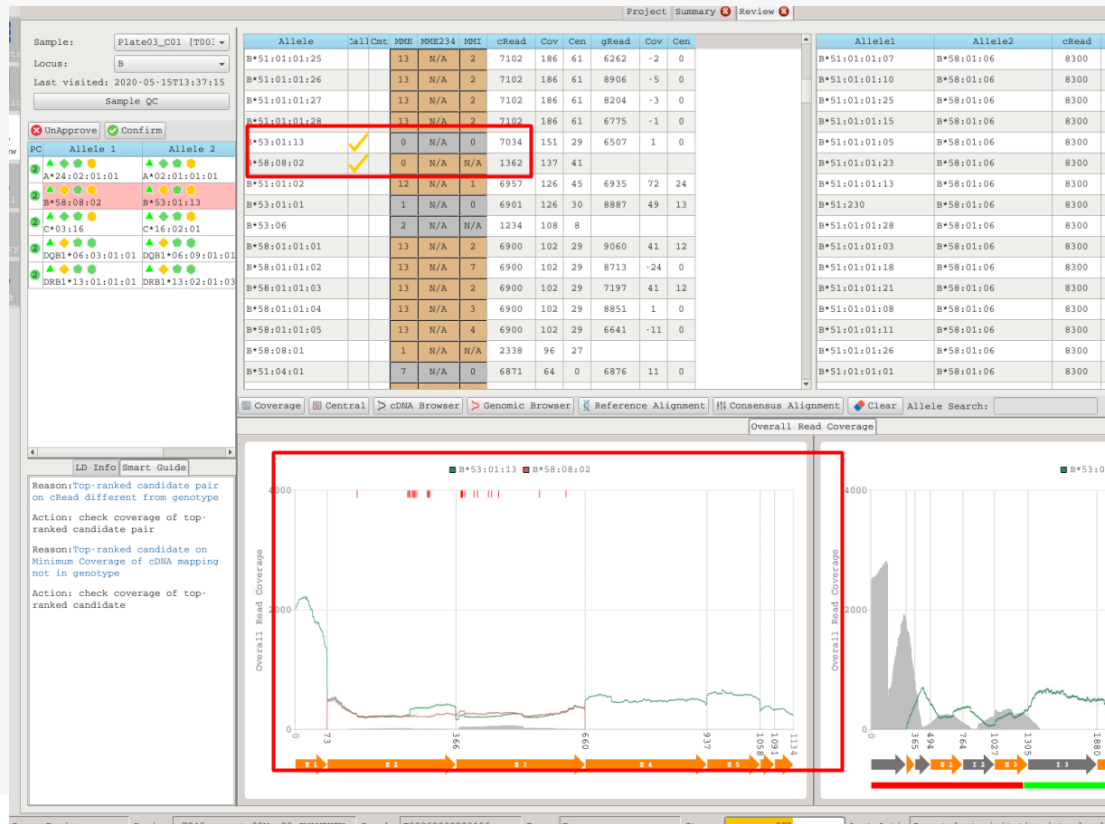
Classic Phasing: Based on statistical analysis of all the reads, the Classic algorithm determines for each position whether it is homozygous or heterozygous. When heterozygous, the Classic algorithm determines for each heterozygous position its phase compared to the neighboring heterozygous positions.

Cluster Phasing: The Cluster phasing algorithm separates the reads into two groups by identifying groups of reads that are likely to originate from the same allele.

Dynamic Phasing: The Dynamic phasing algorithm separates the reads by identifying clusters of reads that are likely to originate from the same allele. Potential overlap between different clusters is taken along in the analysis. In addition, this algorithm incorporates the basecall QV values and can detect local allele imbalances.

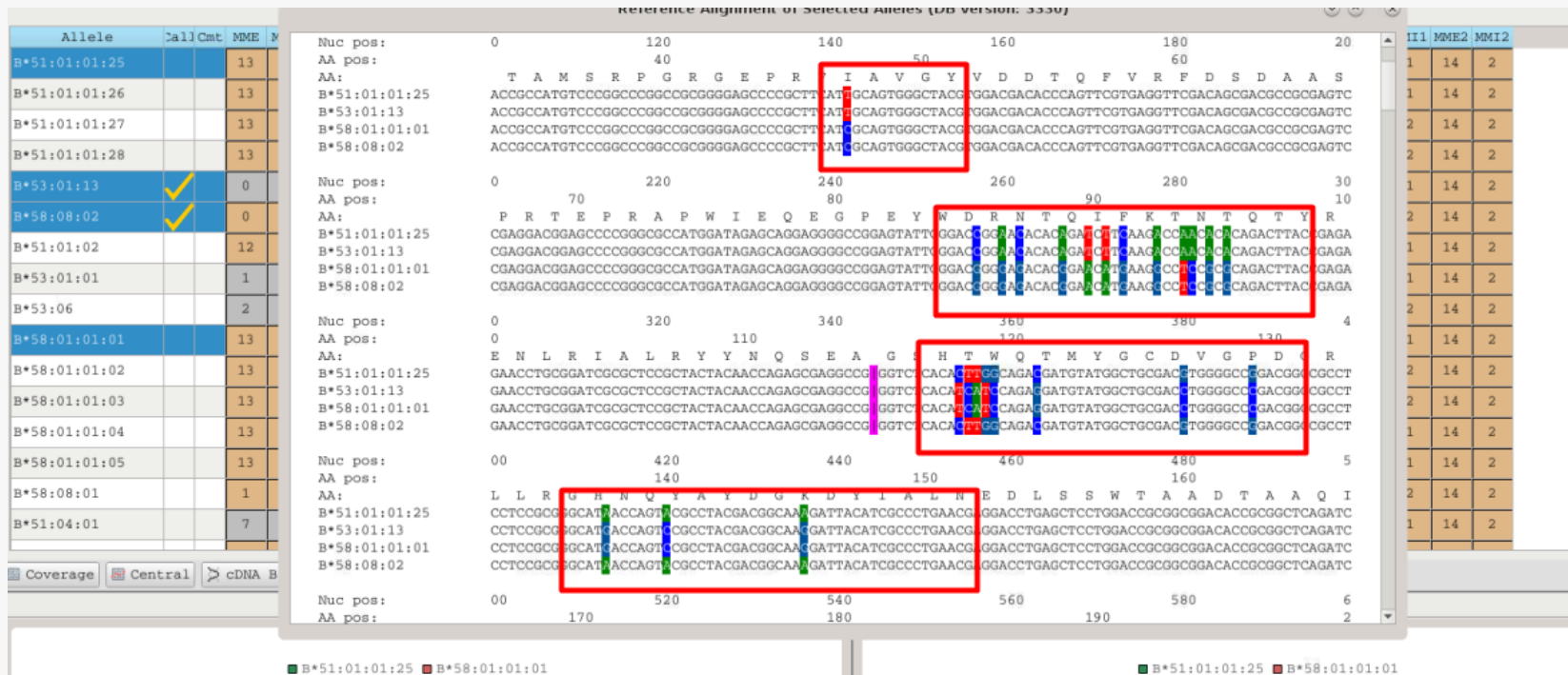
Ambiguity

B*53:01:13/58:08:02- B*51:01:01:01/58:01:01:01



Ambiguity

B*53:01:13/58:08:02- B*51:01:01:01/58:01:01:01



Yeni Nesil Dizileme İş Akışı – Veri Analizi

Etken Faktörleri

- Kısa fragment
- PCR Artifact
- Nükleotid dağılımına bağlı gürültü
- Sekans Platformuna bağlı kısıtlamalar
- Örnek barkodlamasında kontaminasyon
- Dizileme kalitesi



TEŞEKKÜRLER

