

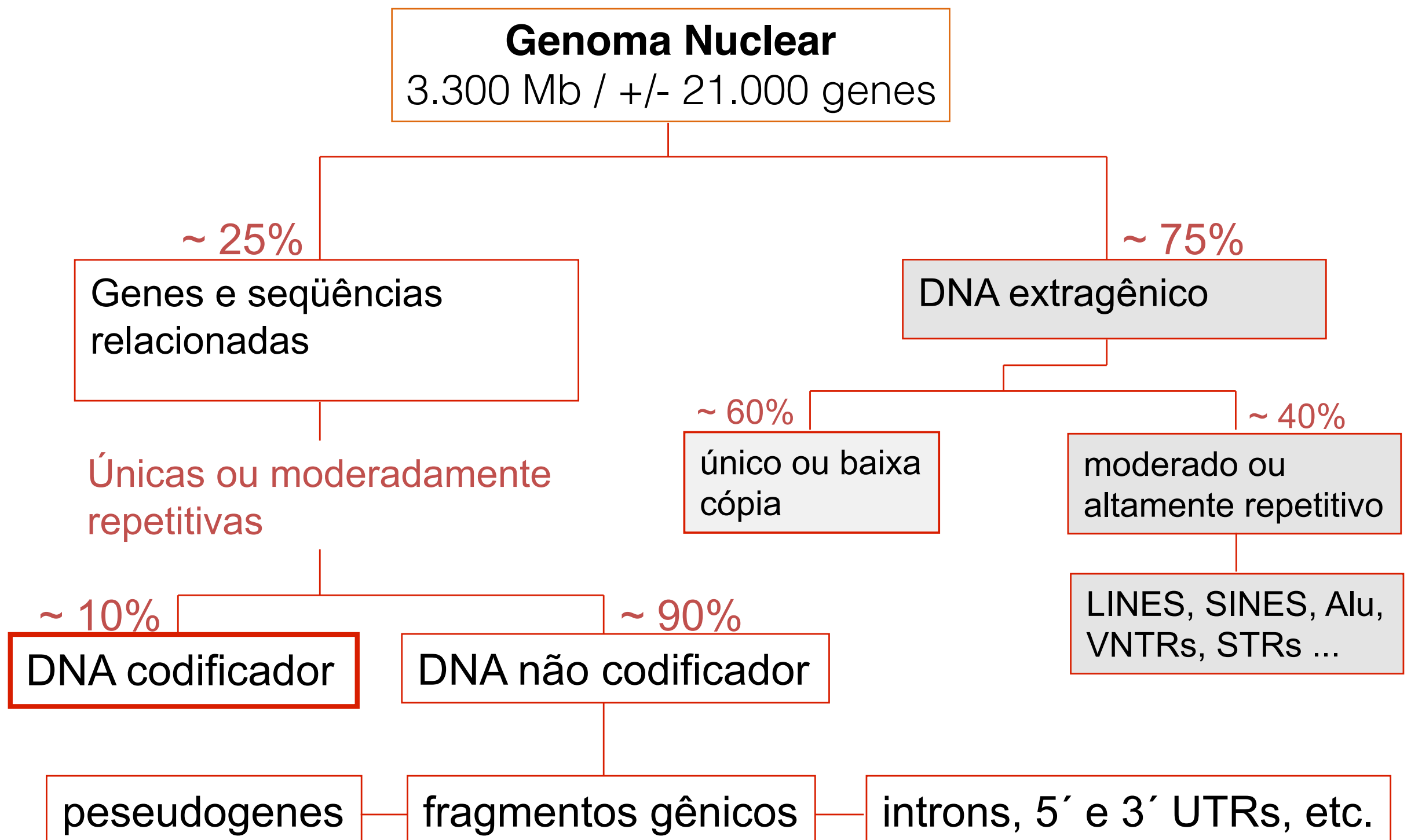
Ciências Biomédicas - FMRP/USP

Disciplina
RCB0300- Biotecnologia III

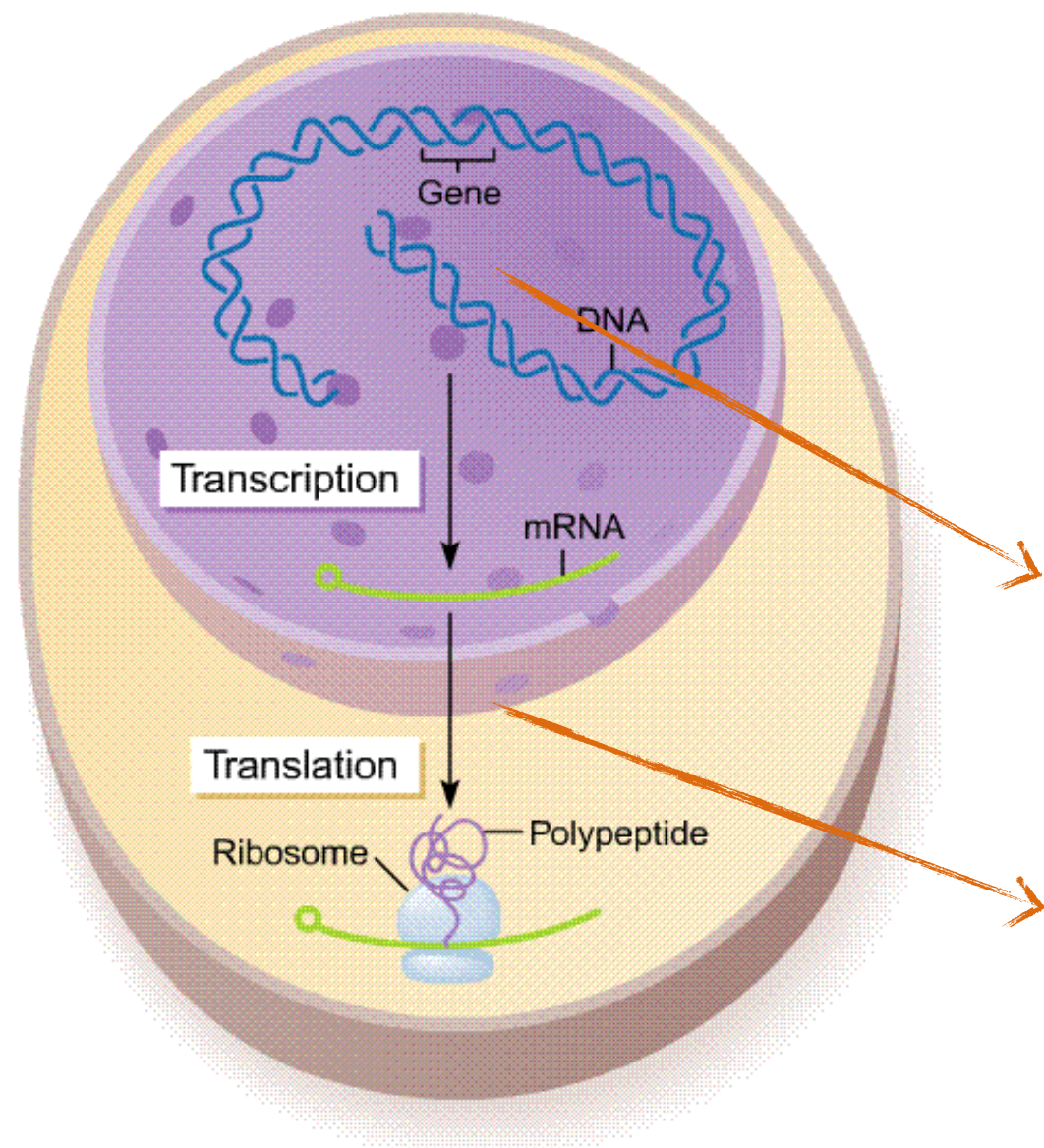
Aula

Princípios de Sequenciamento em escala genômica

Wilson Araújo da Silva Junior
Departamento de Genética da FMRP/USP
<http://rge.fmrp.usp.br>



Dóγμα Central



Transcrição

- Chip-Seq
- Methyl-Seq
- Exome-Seq
- **mRNA-Seq**

Instabilidade
genômica e
regulação da
expressão gênica

Pós-transcrição

- **mRNA-seq**
- miRNA-seq

Regulação da expressão
gênica e de vias gênicas
canônicas e não-canônicas)

Protocols:

DNA

- Whole Genome-Seq
- Exome-Seq

Structural and genomic
alterations/mutation

Epigenomics

- Chip-Seq
- Methyl-Seq

Regulation of Gene Expression

RNA

- mRNA-Seq
- miRNA-Seq

Expression of Genes
(canonical and non-canonical)

Análise de dados em escala genômica

Next Generation Sequencing

1. Definição do problema biológico;

2. Discussão do desenho experimental;

- Tipo de amostra, número amostral, abordagem de NGS, etc.

4. Discussão dos resultados;

5. Conclusão do estudo.

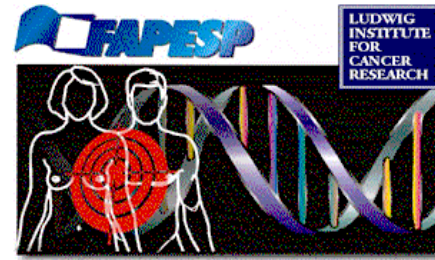
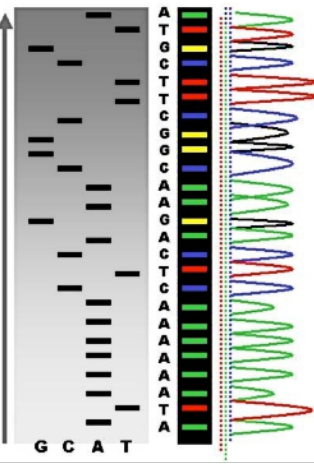


<http://ctcusp.org/>

CISBi-NAP/USP - [Agendamento para suporte em análise de bioinformática](#)



Leroy Hood



Genoma
Humano do Câncer

Francis Collins
Diretor do "National Human Genome Research Institute"



1990

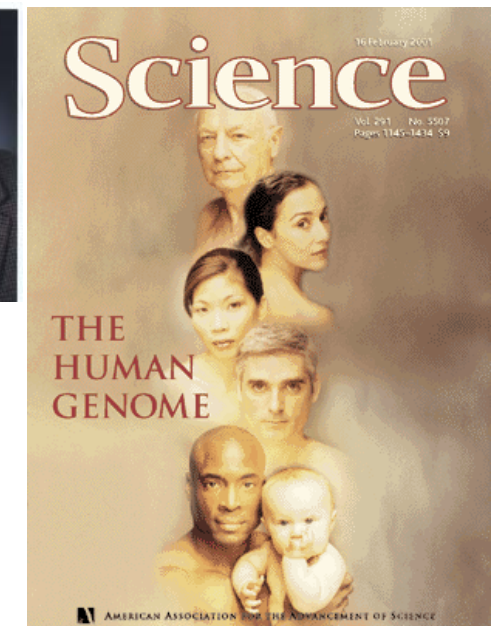
1999

2001

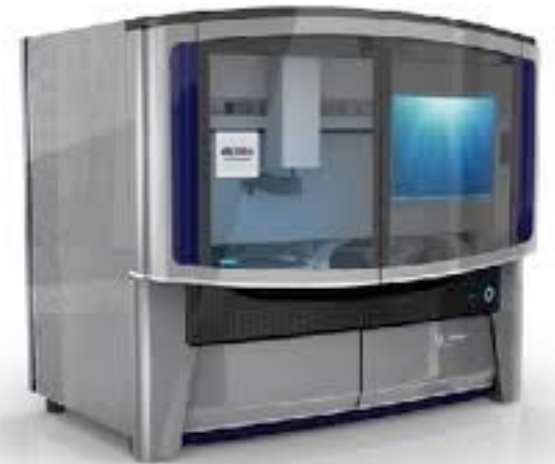
Lançado o Projeto Genoma Humano (HGP)

Esforço de 15 anos coordenado pelo "Department of Energy – DOE" e pelo "National Institutes of Health – NIH" dos USA para o seqüenciamento completo do genoma humano.

Craig Venter
Co-fundador da "Celera Genomics Corporation"



Next Generation Sequencing



2001

2011

Biologia Computacional

Biologia de Sistemas Moleculares



Frederick Sanger
1977

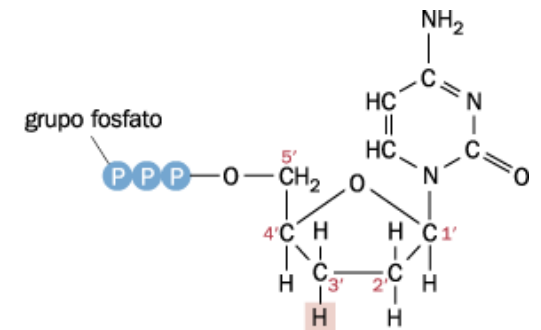
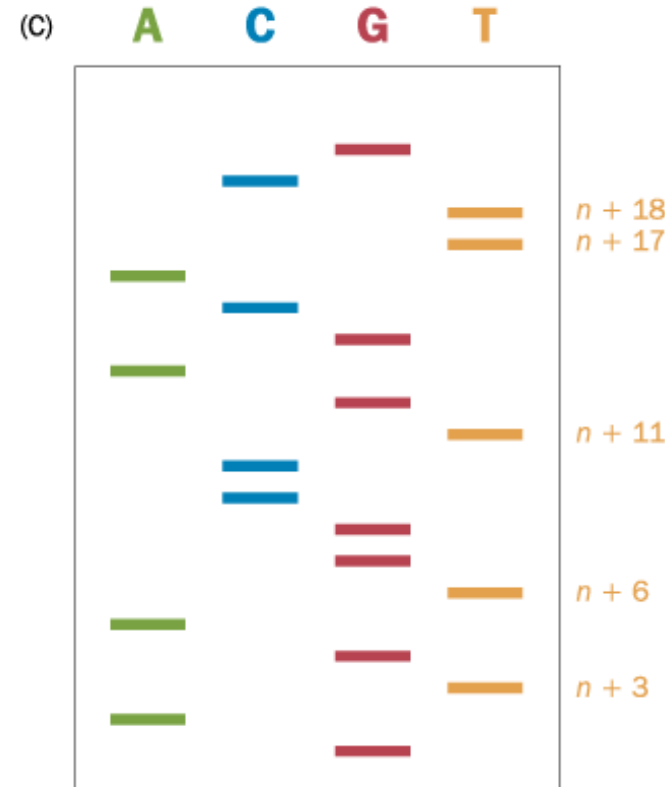
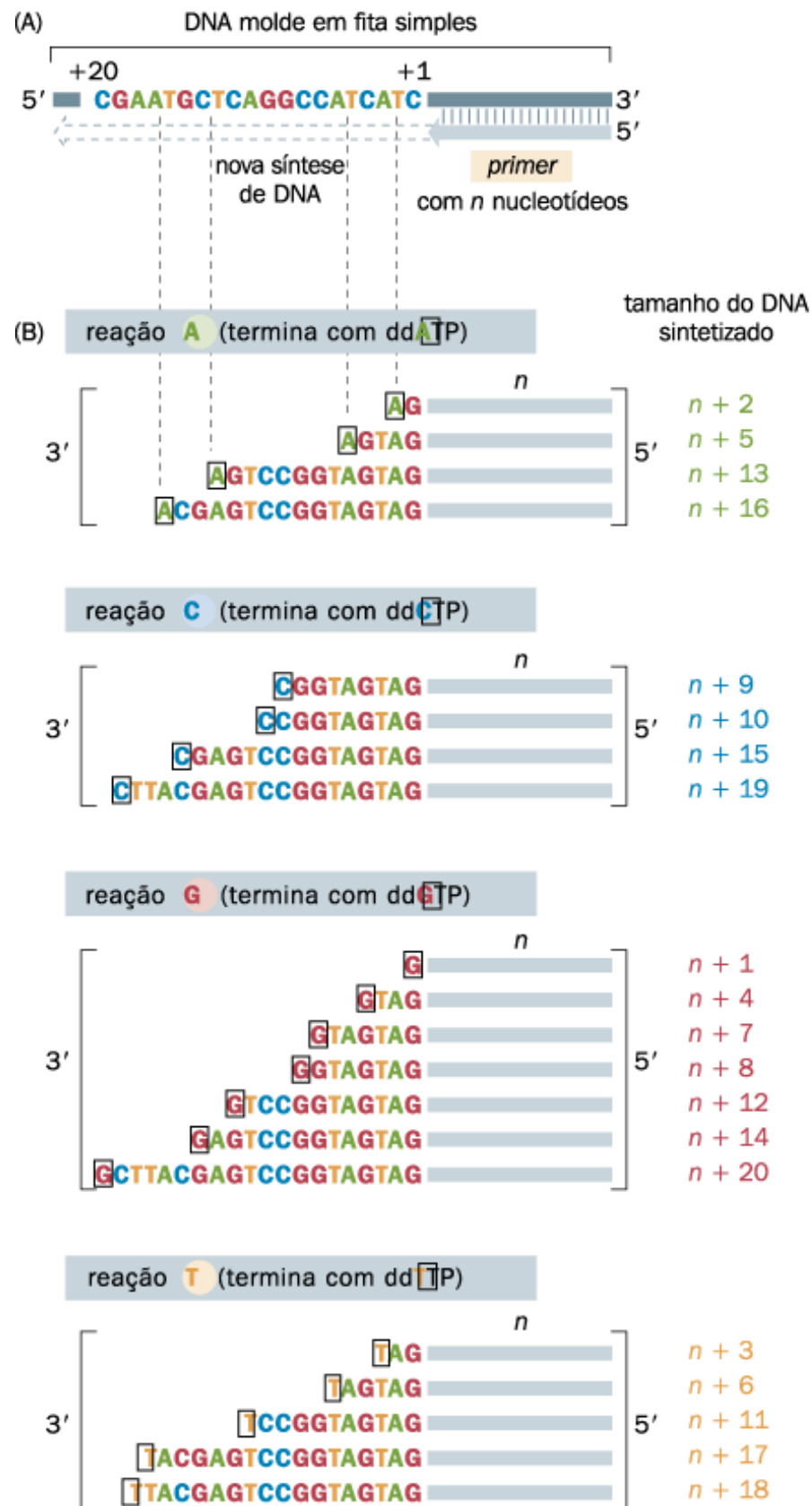
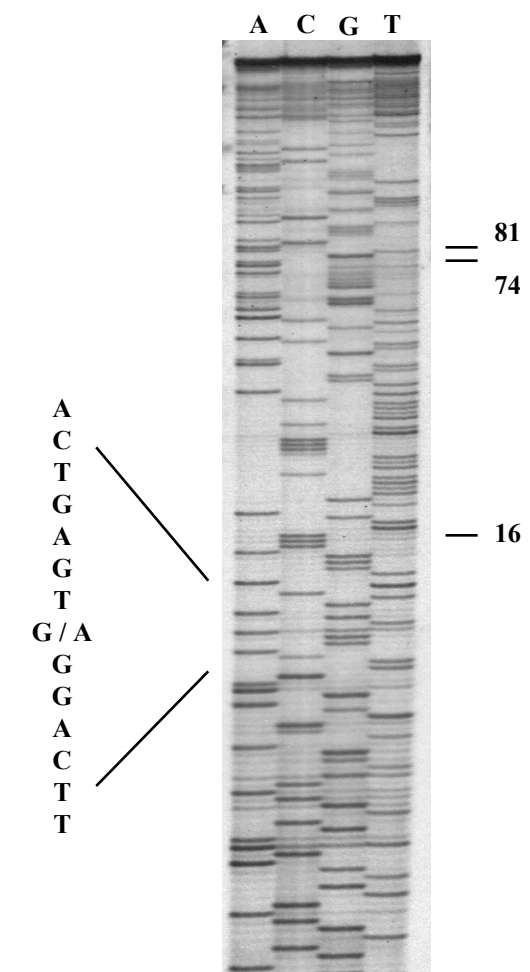
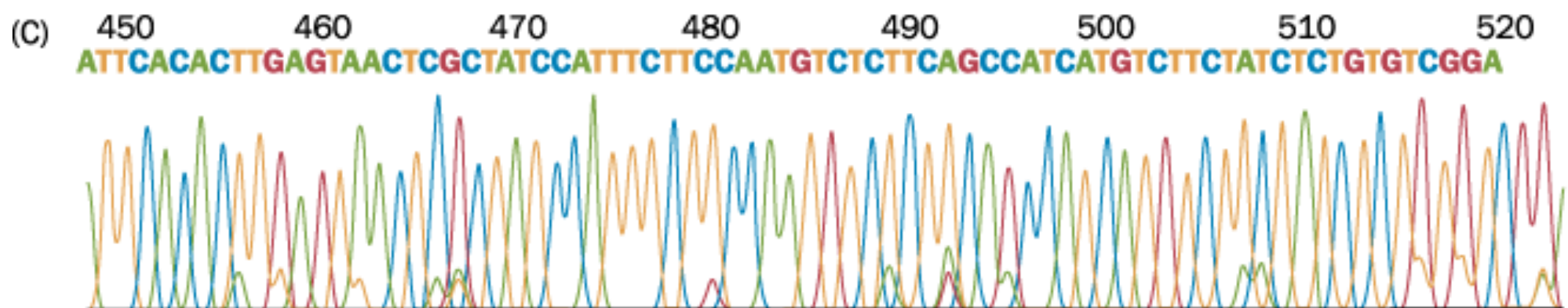
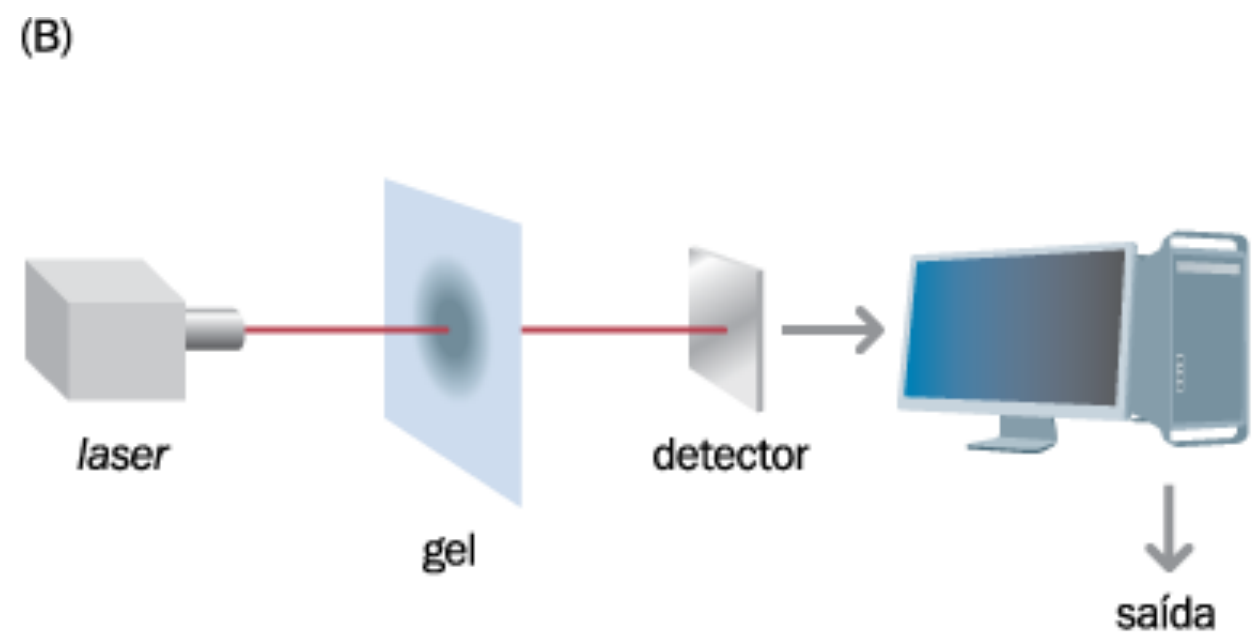
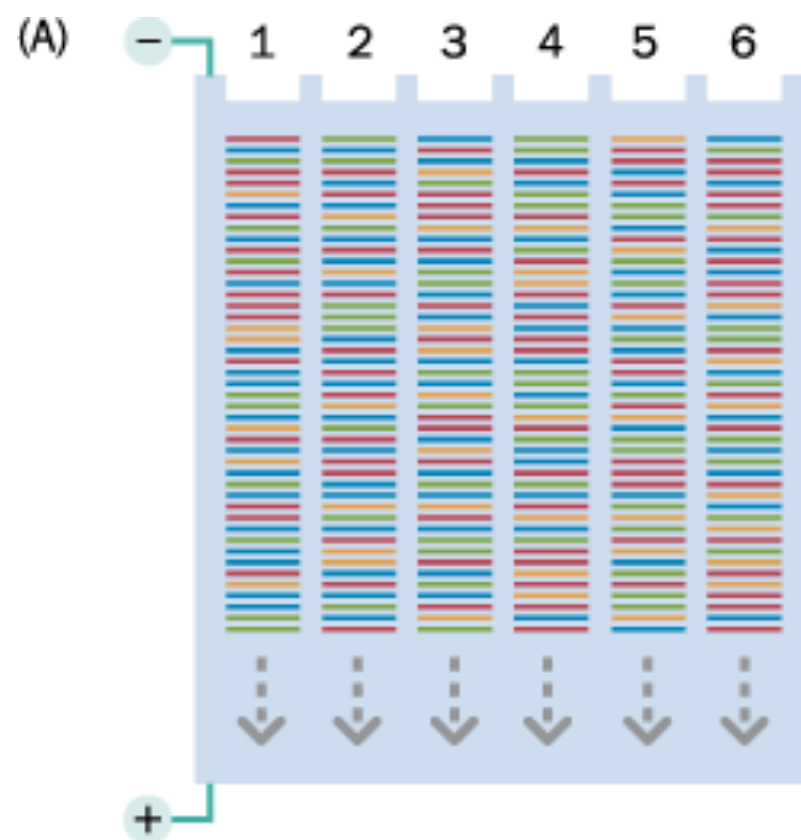


Figura 8.5 Estrutura de um dideoxinucleotídeo: 2', 3'-dideoxi CTP (ddCTP). O grupo hidroxila ancorado ao carbono 3' em nucleotídeos normais é trocado por um átomo de hidrogênio (sombreado).



Hb A/ β^0 -Talassemia (IVS-2 splice junction G-A)



NGS (Next Generation Sequencing)

Etapas:

1. Extração do DNA ou RNA
2. Preparo da biblioteca (fragmentação randômica)
3. Sequenciamento da biblioteca
4. Análise dos dados (Bioinformática)

NGS (Next Generation Sequencing)

1. Extração do DNA ou RNA

- Manual ou automatizada
- Qualidade (pureza e integridade)

Pureza: Contaminação com RNA/proteínas

OD 260nm /280nm Razão = 1,8 – 2,0

Amostras com razão acima de 2,0 tem muito RNA

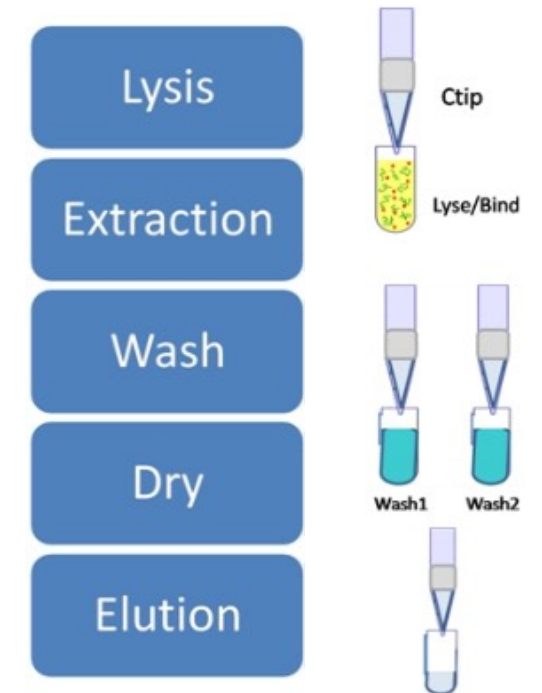
Amostras com razão abaixo de 1,8 tem contaminação com proteína

Contaminação com reagentes

OD 260nm /230nm Razão = 2,0 – 2,2

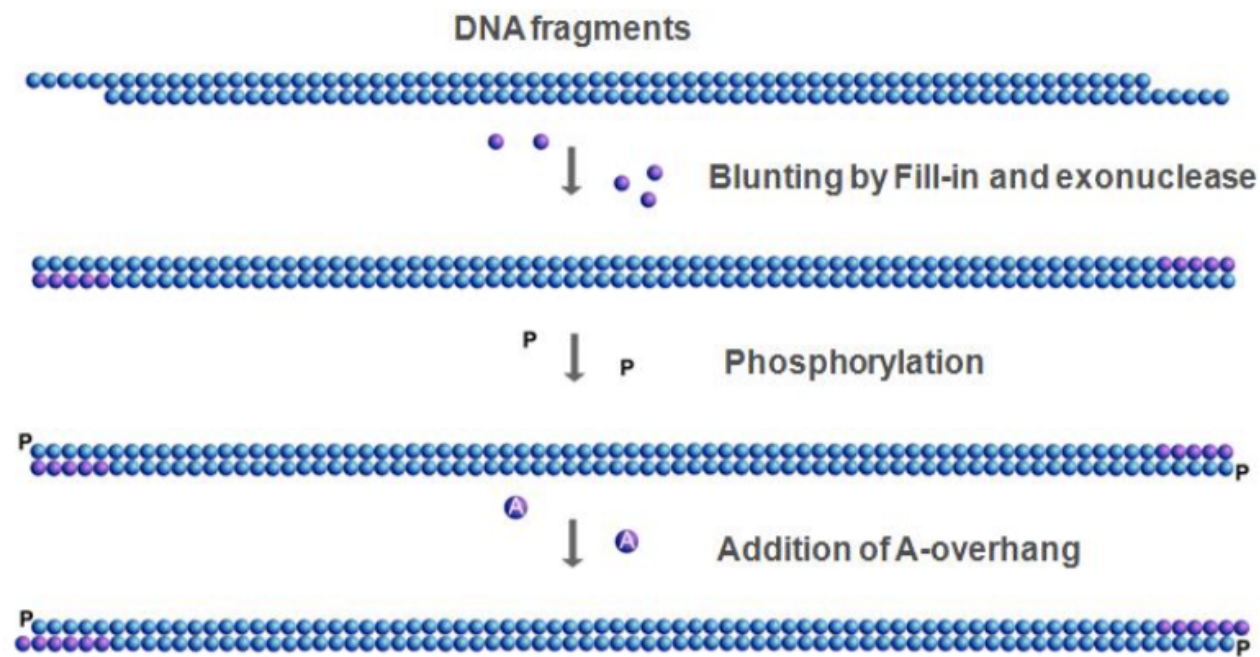
Amostras com razão abaixo do esperado tem contaminação reagentes

Integridade: Fragmentação do DNA ou RNA



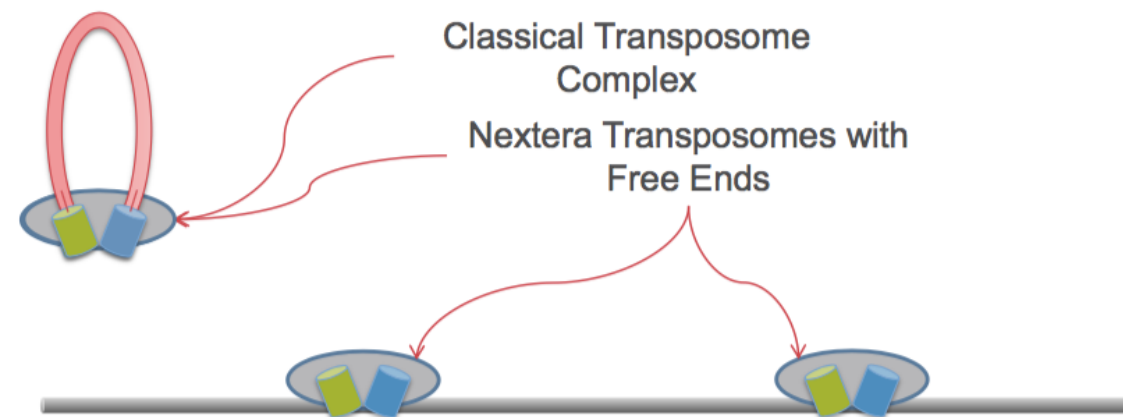
NGS (Next Generation Sequencing)

2. Preparo da biblioteca (fragmentação randômica)



← **Mecânica ou enzimática**

Tagmentação
(Transposase)

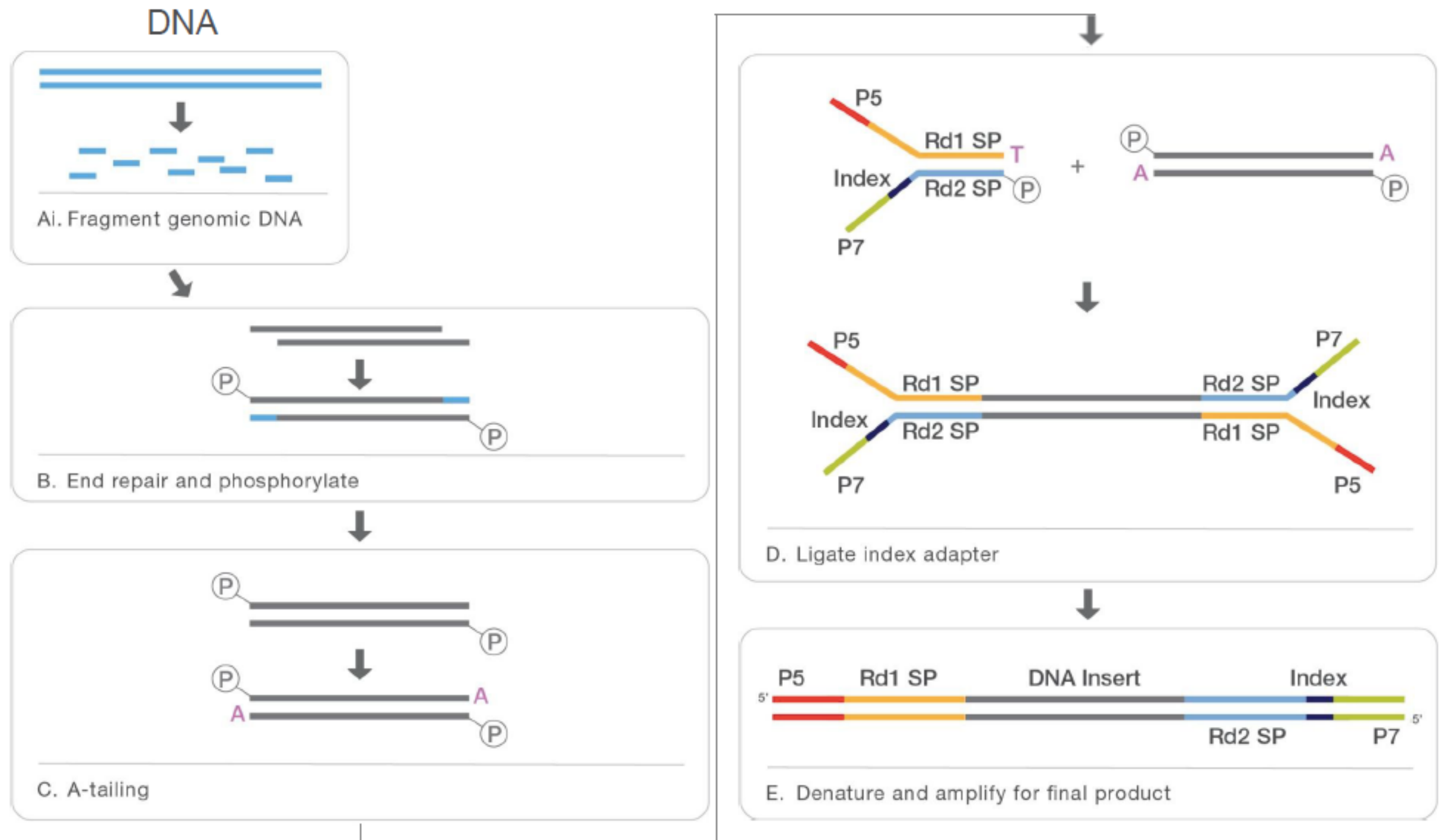


↓ Target DNA is Simultaneously
Fragmented and Tagged



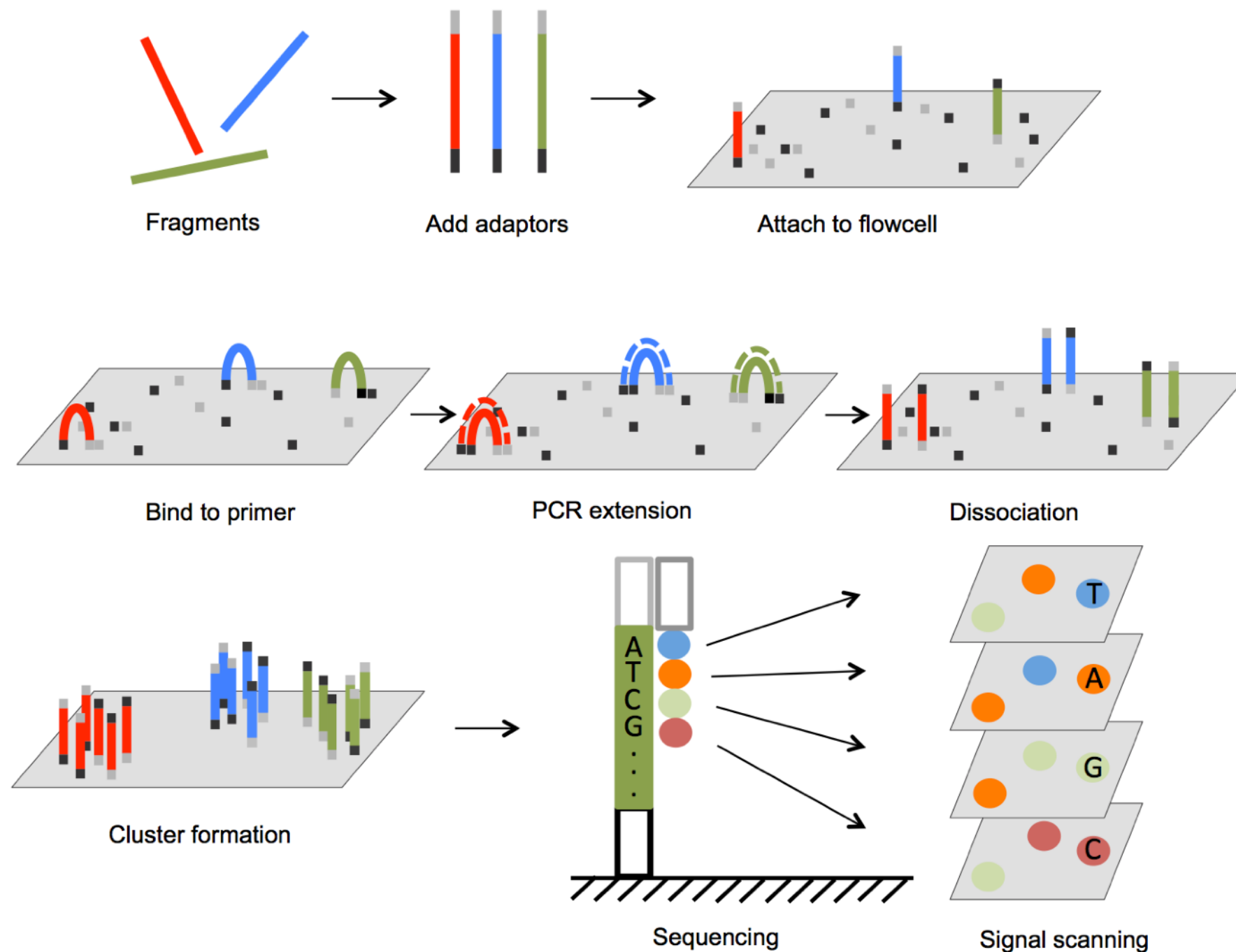
NGS (Next Generation Sequencing)

2. Preparo da biblioteca (fragmentação randômica)



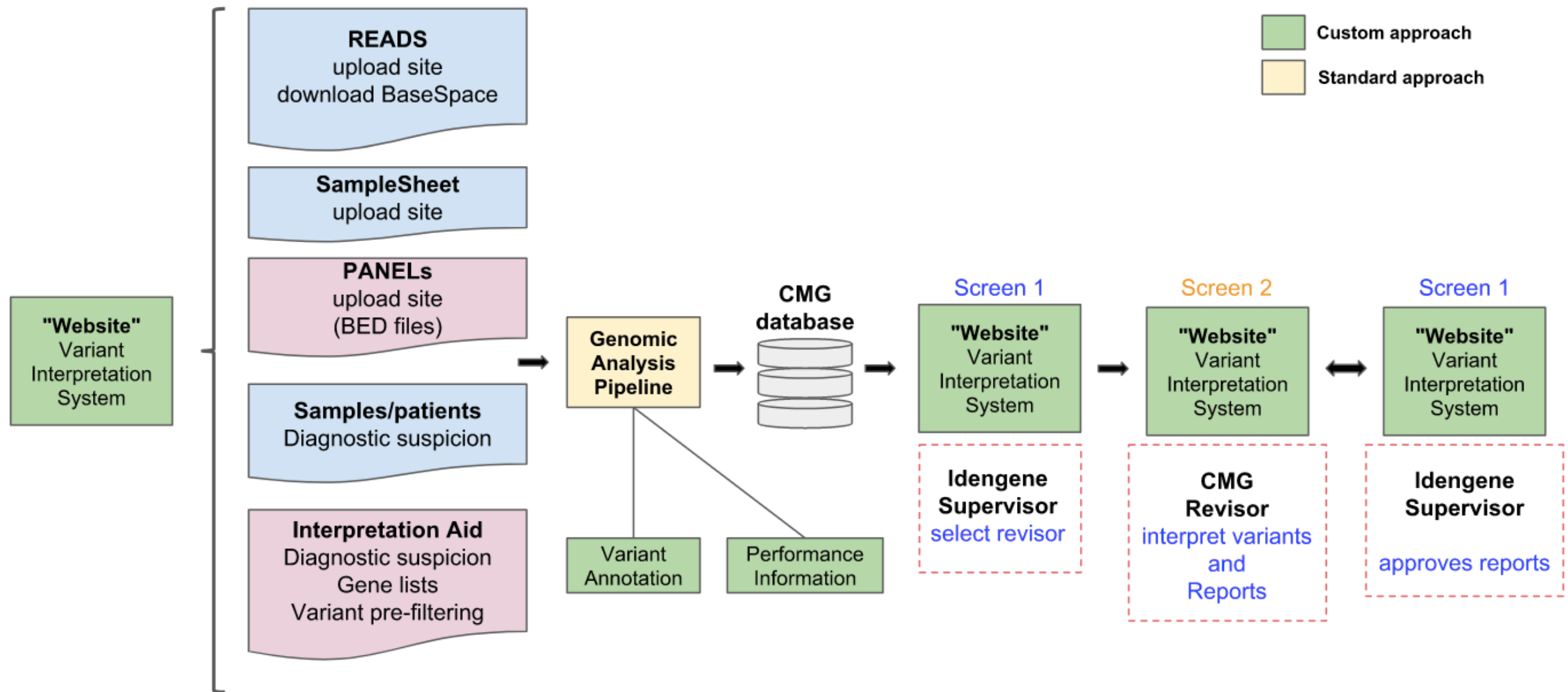
NGS (Next Generation Sequencing)

3. Sequenciamento da biblioteca



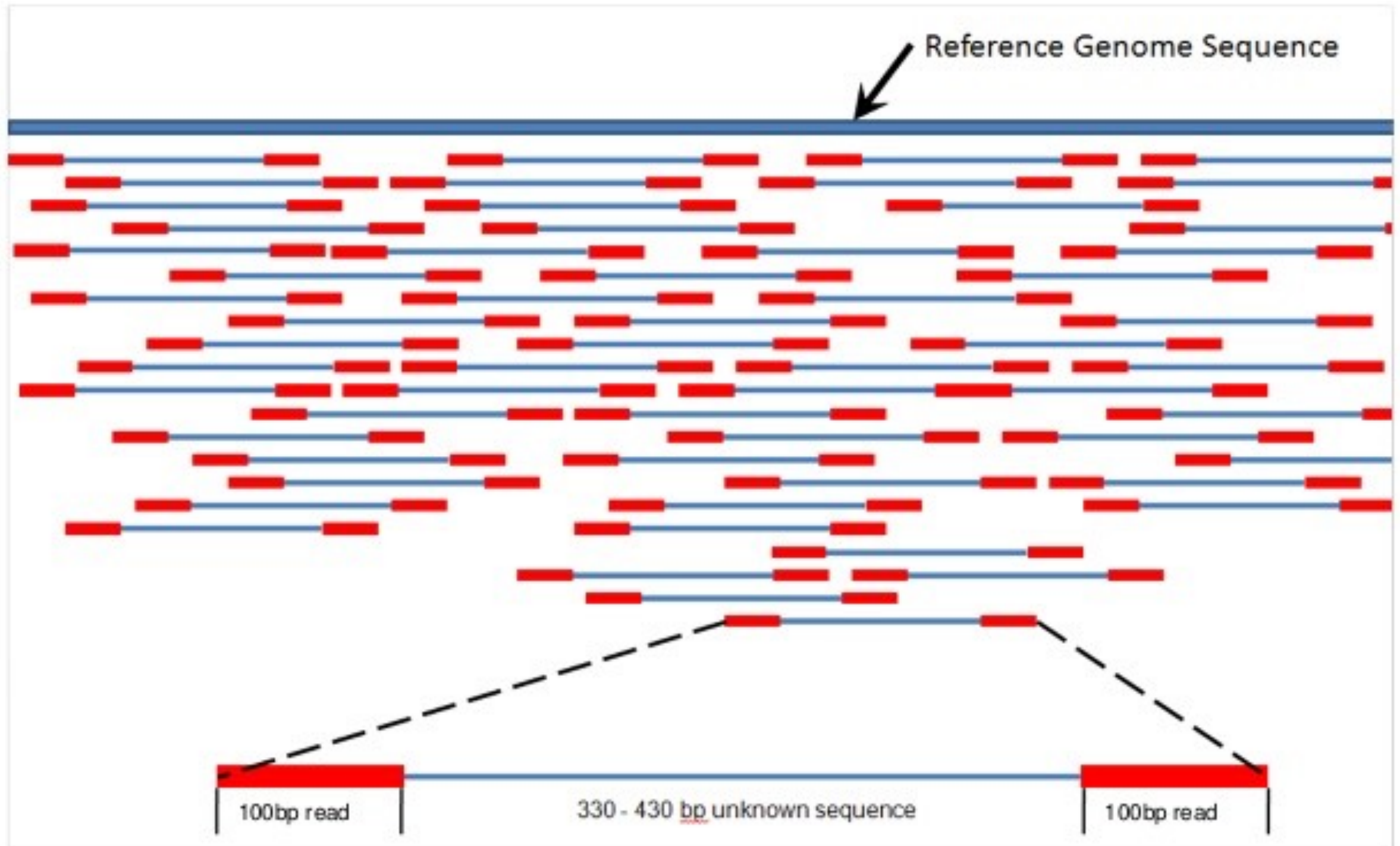
NGS (Next Generation Sequencing)

4. Análise dos dados (Bioinformática)



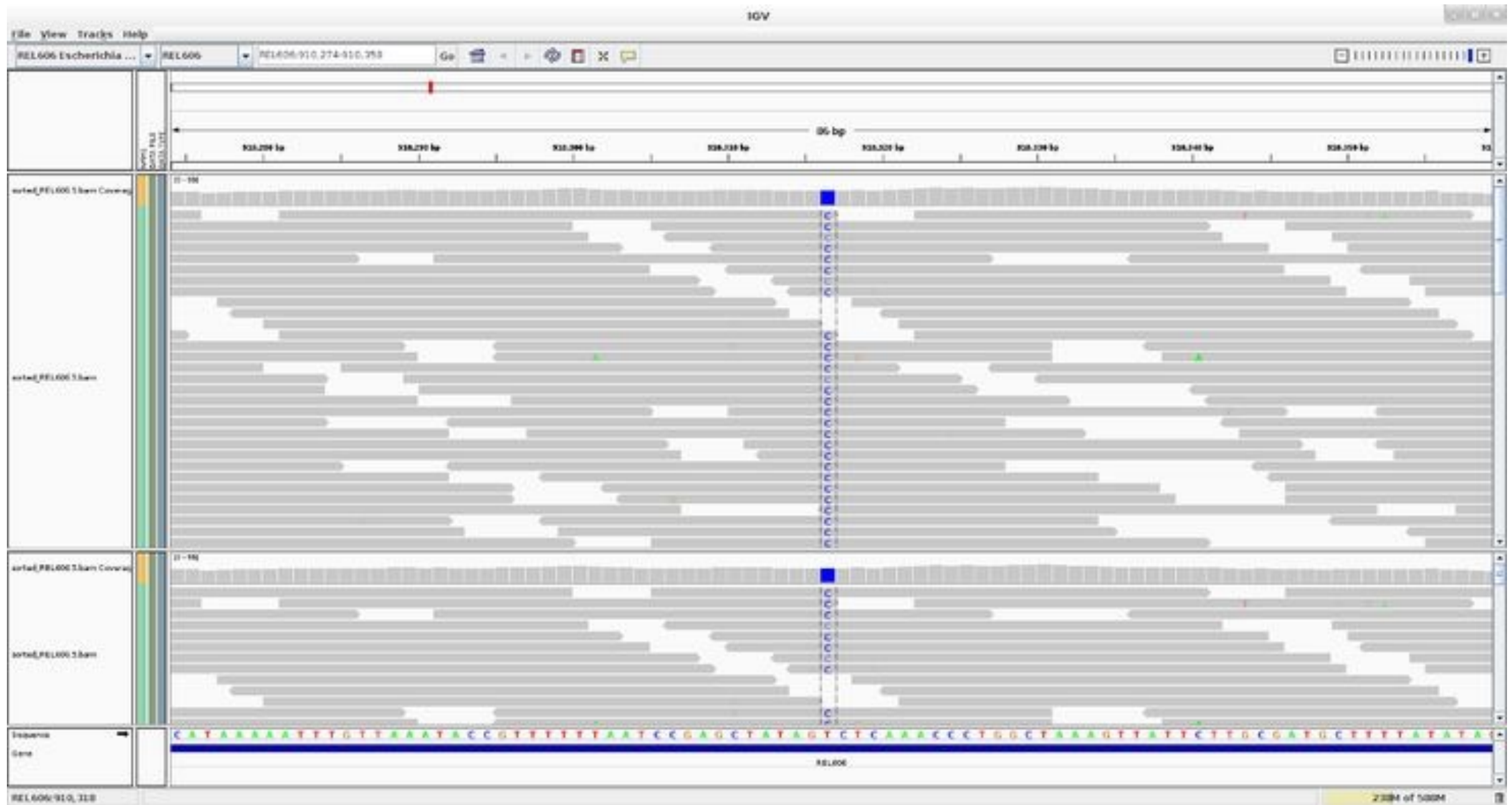
Parâmetros importantes

Cobertura!!!



Parâmetros importantes

Cobertura!!!



Parâmetros importantes

The Lander/Waterman equation

Equação Geral: $C = LN/G$

C: cobertura

L: tamanho da read (sequência)

N: número de reads

G: tamanho do genoma haploide

$$C = (100 \text{ bp}) * (189 \times 10^6) / (3 \times 10^9 \text{ bp}) = 6.3$$



iSeq 100 System



MiniSeq System



MiSeq Series +



NextSeq Series +

Run Time	9–17.5 hours	4–24 hours	4–55 hours	12–30 hours
Maximum Output	1.2 Gb	7.5 Gb	15 Gb	120 Gb
Maximum Reads Per Run	4 million	25 million	25 million [†]	400 million
Maximum Read Length	2 × 150 bp	2 × 150 bp	2 × 300 bp	2 × 150 bp