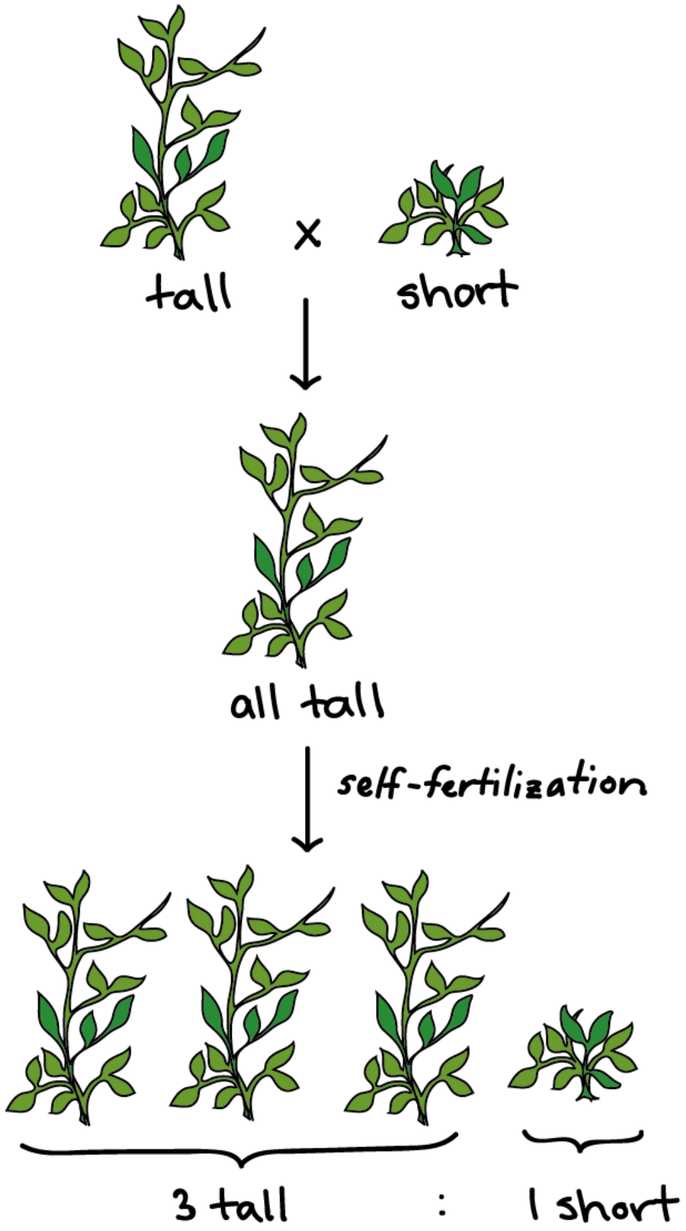
Genômica

Fluxo de informação na célula, estrutura de genomas e estratégias de sequenciamento de DNA

CAROLINA NEGRI
GUSTAVO DO NASCIMENTO
JOÃO ARAUJO

Basic principles of heredity

1866

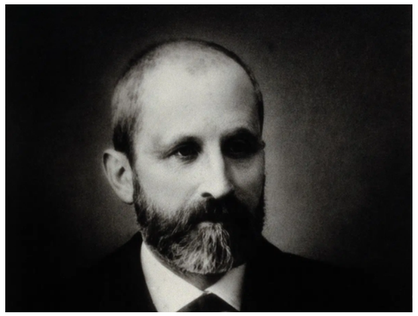


https://www.yourgenome.org/theme/timeline-history-of-genomics/

Basic principles of heredity

1866

1871



Friedrich Miescher identifies the presence of ‘nuclein’

A

Medicinisch-chemische

UNTERSUCHUNGEN.

Aus dem

Laboratorium für angewandte Chemie zu Tübingen

herausgegeben

von

Dr. FELIX HOPPE-SEYLER

o. ö. Professor der angewandten Chemie an der Universität Tübingen.

VIERTES HEFT.

Mit einem Holzschnitt.

BERLIN, 1871.

Verlag von August Hirschwald.

Unter den Linden. 68.

B

XLV.

Ueber die chemische Zusammensetzung der Eiterzellen [1]).

Von Dr. **F. Miescher** aus Basel.

Die Chemie des Eiters ist bis vor Kurzem fast nur von den Gesichtspunkten aus studiert worden, die für die Untersuchung von pathologischen Transsudaten massgebend waren. In neuerer Zeit hat man sich mit der Erforschung der Eigenschaften des Protoplasma auch an die Eiterzellen gewandt. Insbesondre musste sich aber seit den bekannten Untersuchungen über die Herkunft der Eiterzellen der Gedanke aufdrängen, dass hier das nächstliegende Material sei zum Studium dieser Zellenspezies, die als constante Grösse nunmehr an so vielen Orten wird eingeführt werden müssen; ein Material, nicht tadelfrei, mit Vorsicht zu verwerthen, aber das einzige leicht zu beschaffende und desshalb zum vorläufigen Ausgangspunkt geeignet.

In diesem Sinne habe ich versucht, über die eigentlich gewebsbildenden Stoffe in den Eiterzellen zu einiger Orientierung zu gelangen. Die ganze Reihe der Extractivstoffe, in sofern sie ihrer Menge und Beschaffenheit nach nicht als wesentliche Gewebsbildner zu betrachten sind, habe ich bei Seite gelassen. Das Material zur Untersuchung wurde mir durch dankenswerthe Vermittlung der Herren Assistenzärzte Dr. Beyer und Dr. Koch aus der Tübinger chirurgischen Klinik geliefert. Die Verbände, weitaus überwiegend von Operationswunden herrührend, wurden gesammelt, täglich auf das Laboratorium

1) Die Untersuchungen, welche Hr. Miescher in dieser Abhandlung schildert, sind im Tübinger Schlosslaboratorium von Herbst 1868 bis Herbst 1869 ausgeführt und mir kurz darauf zur Veröffentlichung in diesem Hefte übergeben, dessen Erscheinen durch mehre unvorhergesehene Umstände sehr verzögert ist.
Hoppe-Seyler.

Hoppe-Seyler, med. chem. Unters. 30

C

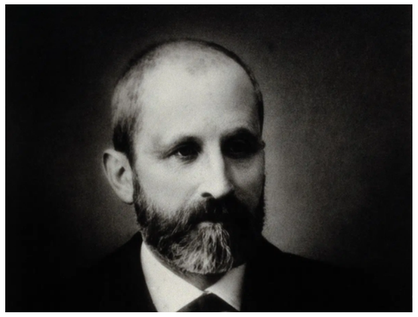
Darstellung, welche den N-gehalt niederdrücken musste. — Auf Grund blos qualitativer Versuche würde man an eine Verbindung von Lecithin mit einem Eiweisskörper oder Eiweisskörperähnling denken, etwa wie man sich das Vitellin oder Ichthin vorgestellt hat. Aber 5,8 % PO_4 und 14 % N in einer und derselben Substanz lassen sofort diese Annahme dahinfallen. Wir haben vielmehr hier Körper sui generis, mit keiner jetzt bekannten Gruppe vergleichbar. Etwaige vielleicht berechtigte Zweifel an der vollkommenen Reinheit meines Präparats ändern an dieser Thatsache nichts. Dass der Phosphor wirklich an die organische Substanz gebunden ist, davon habe ich mich an zwei Versuchen überzeugt, der eine an 0,28, der andere an 0,38 gr. trockener Sub-

https://www.yourgenome.org/theme/timeline-history-of-genomics/

1866

Basic principles of heredity

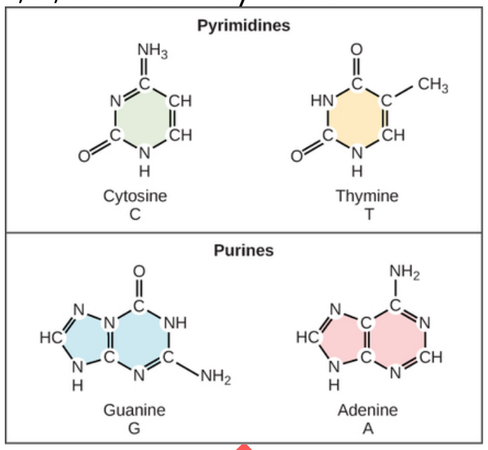
1871



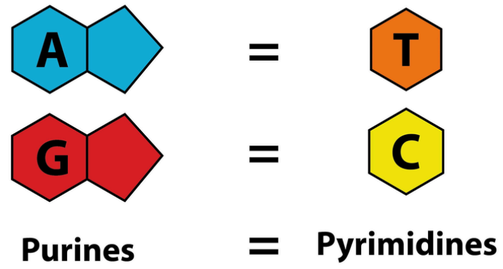
Friedrich Miescher identifies the presence of ‘nuclein’

1910

Discovery of the five DNA bases – A, T, C, G and U – by Albrecht Kossel



1947



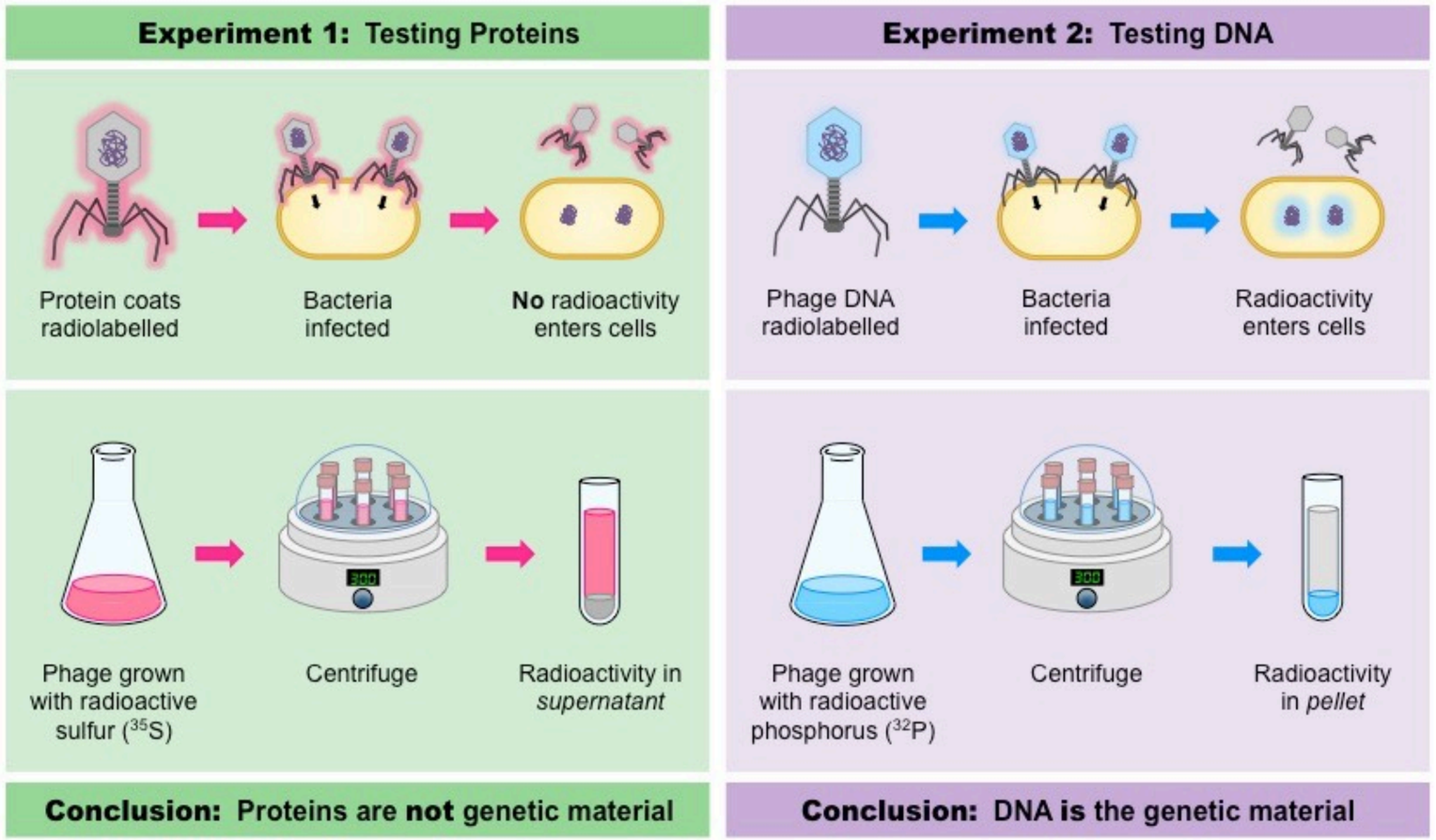
Chargaff’s rule

https://www.yourgenome.org/theme/timeline-history-of-genomics/

DNA, rather than protein, is proven to carry our genetic information – in the Hershey-Chase experiments

Courtesy of Cold Spring Harbor Laboratory Archives. Noncommercial, educational use only.

1950



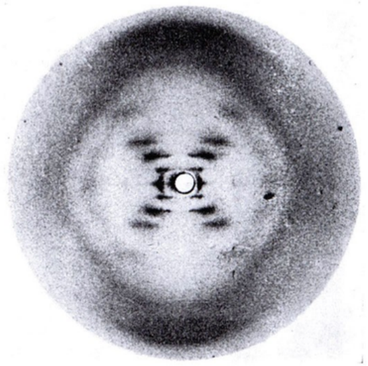
https://www.yourgenome.org/theme/timeline-history-of-genomics/

DNA, rather than protein, is proven to carry our genetic information – in the Hershey-Chase experiments

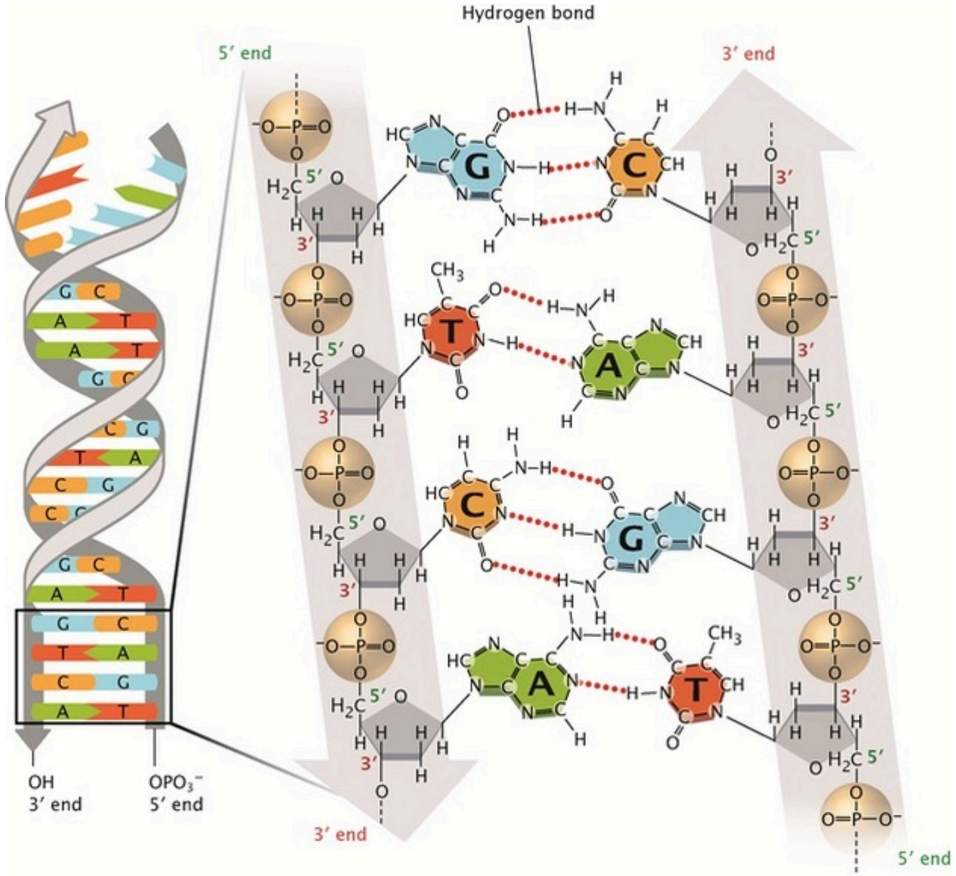
Courtesy of Cold Spring Harbor Laboratory Archives. Noncommercial, educational use only.

1950

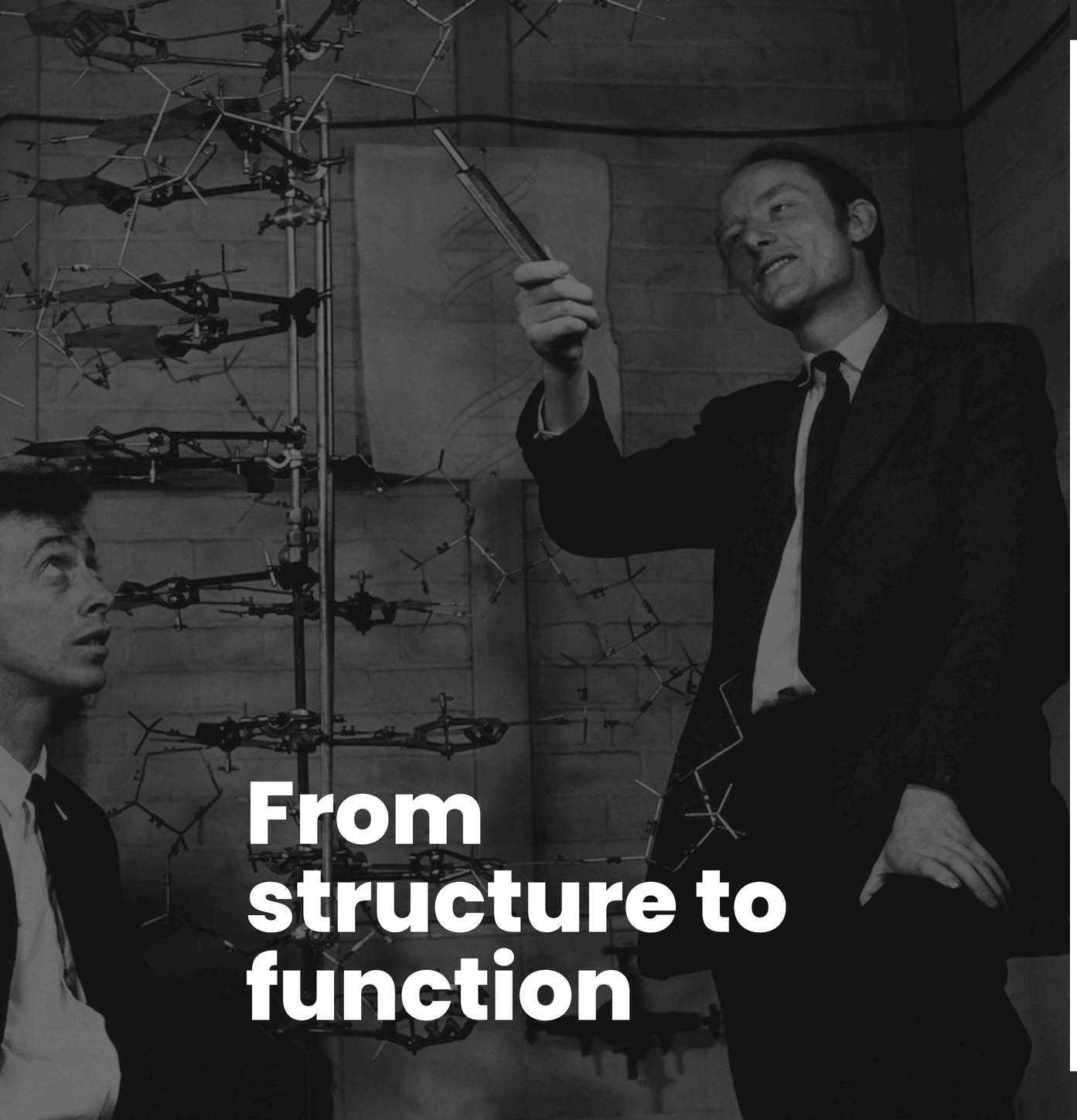
1953



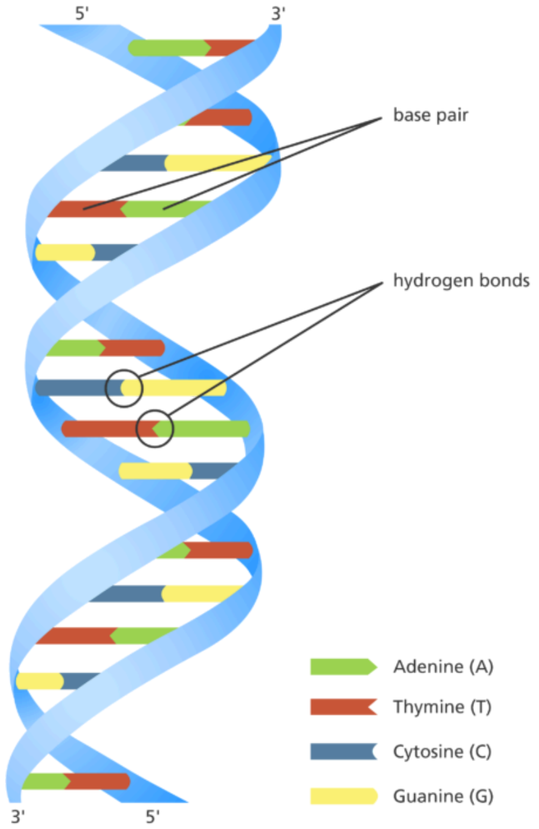
The discovery of DNA's double helix structure by James Watson, Francis Crick, Rosalind Franklin and Maurice Wilkins

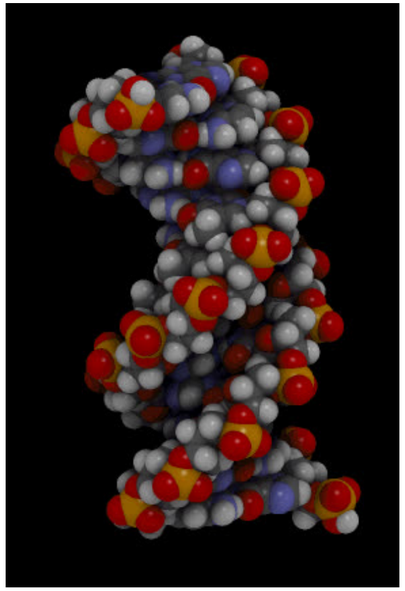


https://www.yourgenome.org/theme/timeline-history-of-genomics/



From structure to function





Feature | Published: 23 January 2003

The double helix and the 'wronged heroine'

Brenda Maddox

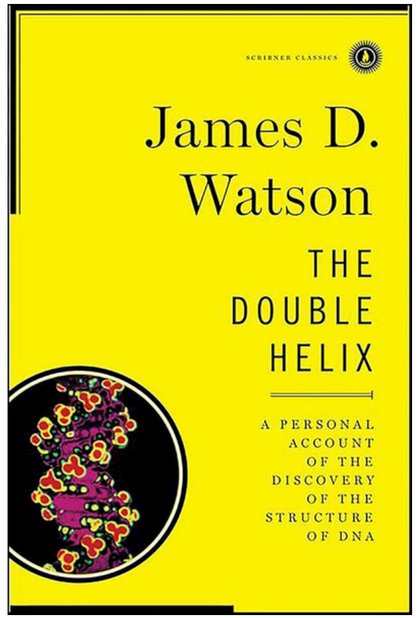
Nature **421**, 407–408 (2003) | Cite this article

COMMENT | 25 April 2023

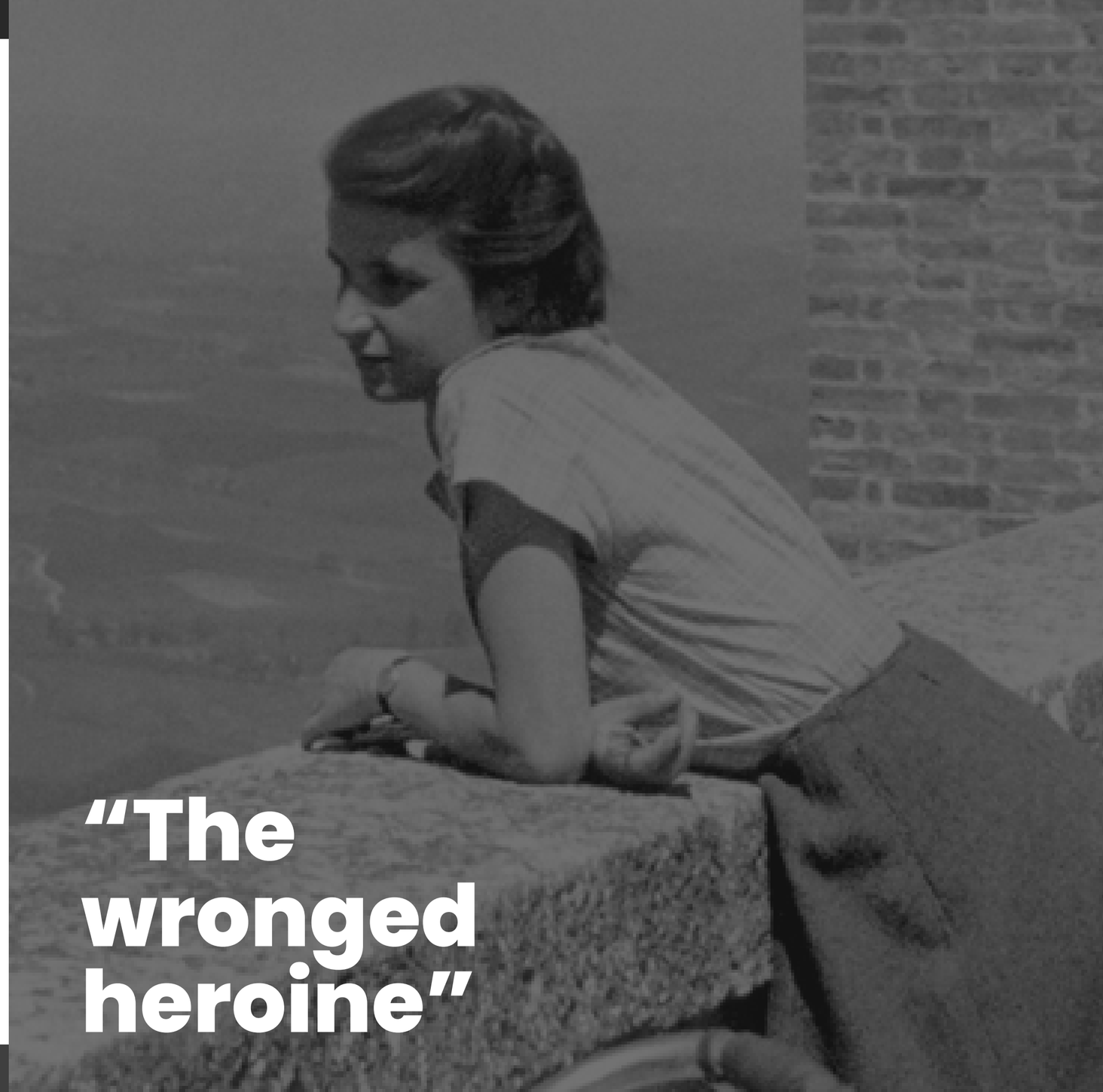
What Rosalind Franklin truly contributed to the discovery of DNA's structure

Franklin was no victim in how the DNA double helix was solved. An overlooked letter and an unpublished news article, both written in 1953, reveal that she was an equal player.

By Matthew Cobb & Nathaniel Comfort



"Rosy, of course, did not directly give us her data. For that matter, no one at King's realized they were in our hands."



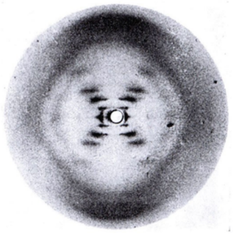
"The wronged heroine"

DNA, rather than protein, is proven to carry our genetic information – in the Hershey-Chase experiments

Courtesy of Cold Spring Harbor Laboratory Archives. Noncommercial, educational use only.

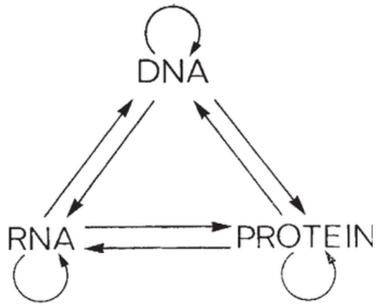
1950

1953



The discovery of DNA's double helix structure by James Watson, Francis Crick and Rosalind Franklin

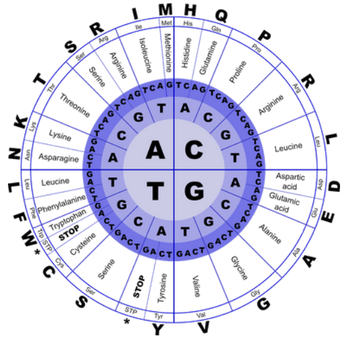
Central Dogma of Molecular Biology - by Francis Crick



1958

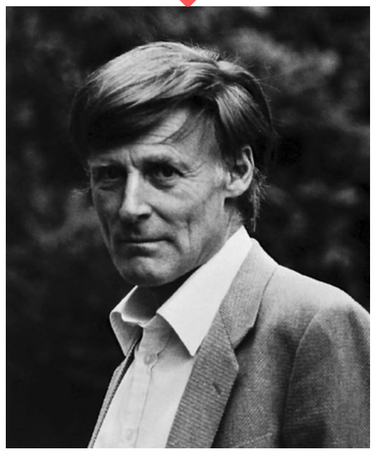
https://www.yourgenome.org/theme/timeline-history-of-genomics/

Cracking the code for life

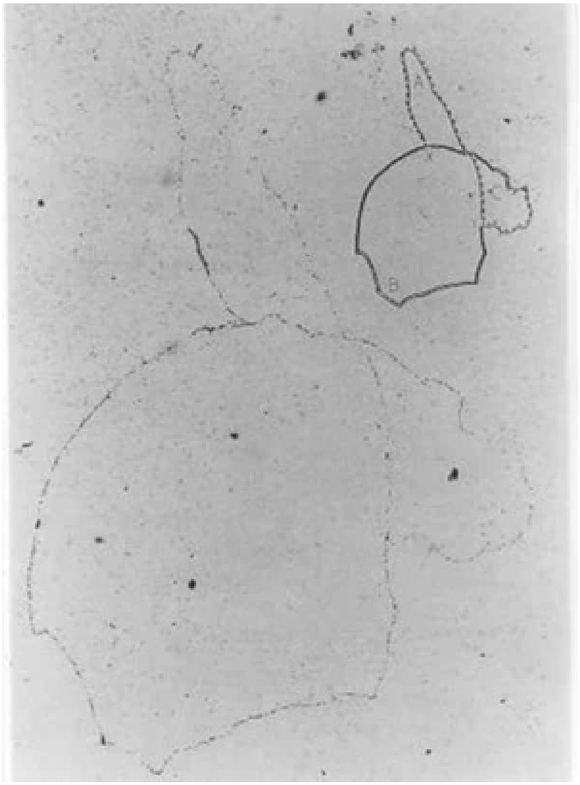


1961

1963



Circular DNA of prokaryotes
– by John Cairns



https://www.yourgenome.org/theme/timeline-history-of-genomics/

Técnicas de Sequencimento de DNA

Shendure, J., Balasubramanian, S., Church, G. et al. DNA sequencing at 40: past, present and future. Nature 550, 345–353 (2017). https://doi.org/10.1038/nature24286

24 bases, This took two years: one base per month

The Nucleotide Sequence of the *lac* Operator

(regulation/protein–nucleic acid interaction/DNA–RNA sequencing/oligonucleotide priming)

WALTER GILBERT AND ALLAN MAXAM

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Communicated by J. D. Watson, August 9, 1973

1968

1973

1977

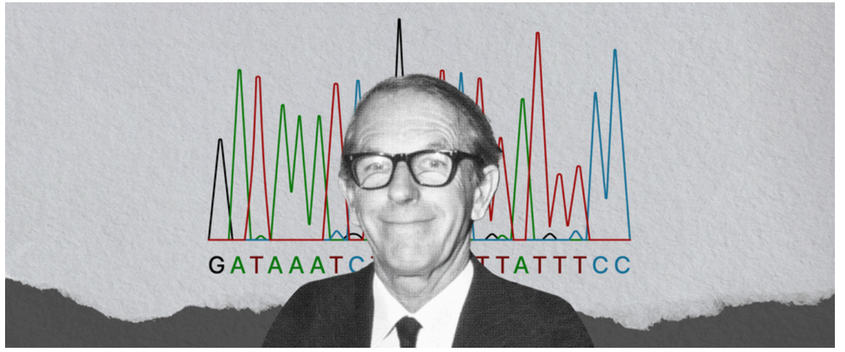
Structure and Base Sequence in the Cohesive Ends of Bacteriophage Lambda DNA

RAY WU AND A. D. KAISER

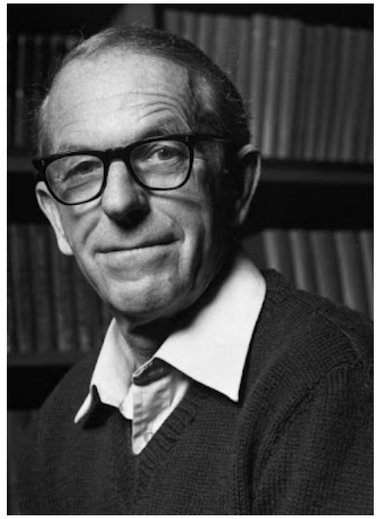
Section of Biochemistry and Molecular Biology, Cornell University Ithaca, New York, and Department of Biochemistry, Stanford University School of Medicine, Palo Alto, California, U.S.A.

(*Received 4 March 1968, and in revised form 6 May 1968*)

12 bases of the cohesive ends of bacteriophage lambda



1ª Geração de Sequenciamento



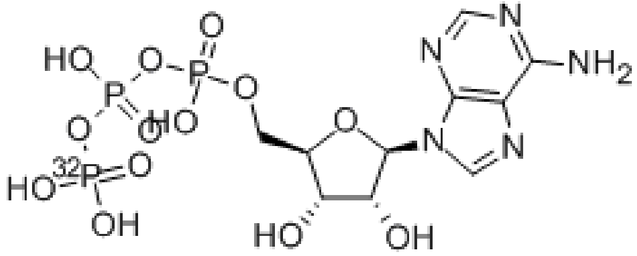
A new method for sequencing DNA

(DNA chemistry/dimethyl sulfate cleavage/hydrazine/piperidine)

ALLAN M. MAXAM AND WALTER GILBERT

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

Contributed by Walter Gilbert, December 9, 1976



Chemical cleavage

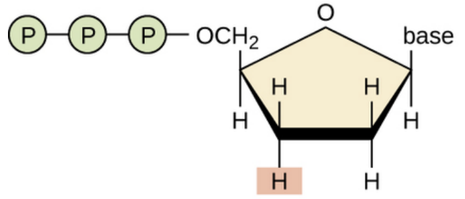
DNA sequencing with chain-terminating inhibitors

(DNA polymerase/nucleotide sequences/bacteriophage ϕX174)

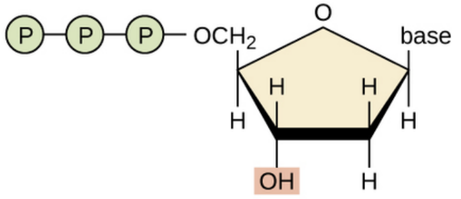
F. SANGER, S. NICKLEN, AND A. R. COULSON

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

Contributed by F. Sanger, October 3, 1977



dideoxynucleotide (ddNTP)



deoxynucleotide (dNTP)

Amplification

A new method for sequencing DNA

(DNA chemistry/dimethyl sulfate cleavage/hydrazine/piperidine)

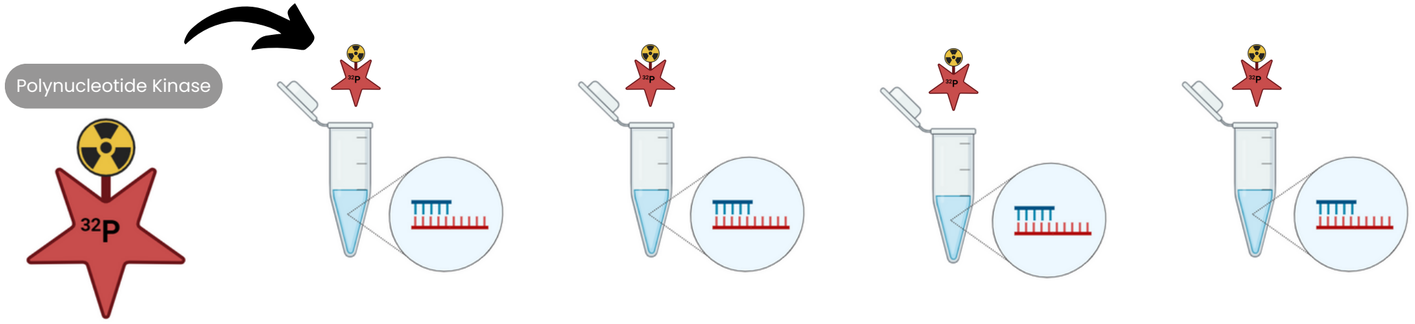
ALLAN M. MAXAM AND WALTER GILBERT

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138

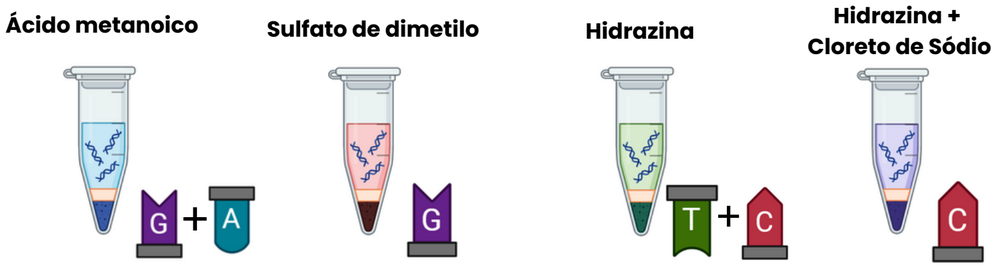
Contributed by Walter Gilbert, December 9, 1976

01. DNA sample

02. Denaturation of DNA strands

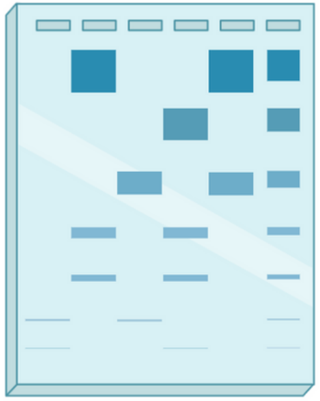


03. Addition of 32P at the 5' end

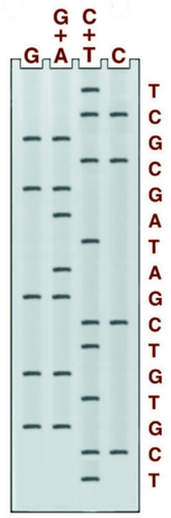


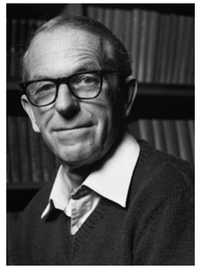
04. Chemical cleavage

05. Polyacrylamide gel



06. DNA Sequecing





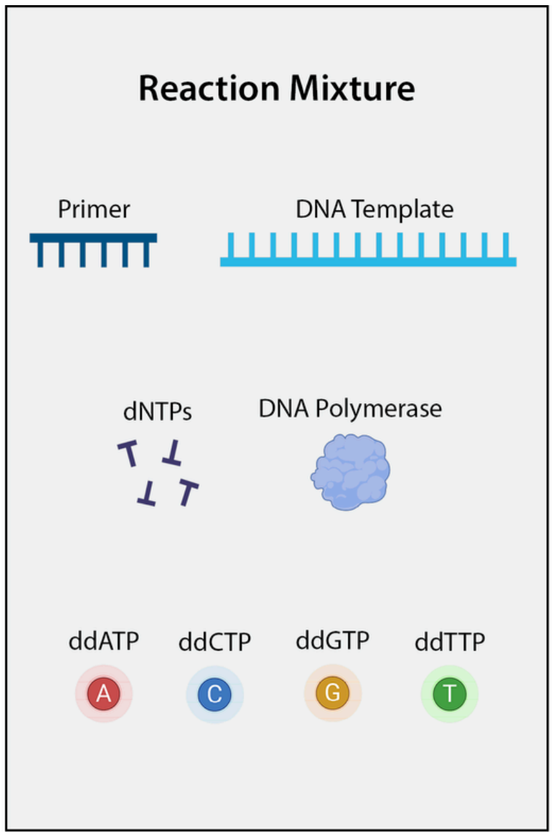
DNA sequencing with chain-terminating inhibitors

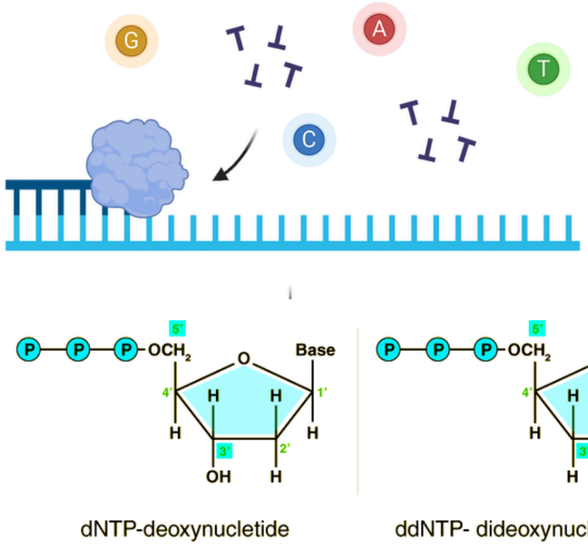
(DNA polymerase/nucleotide sequences/bacteriophage ϕX174)

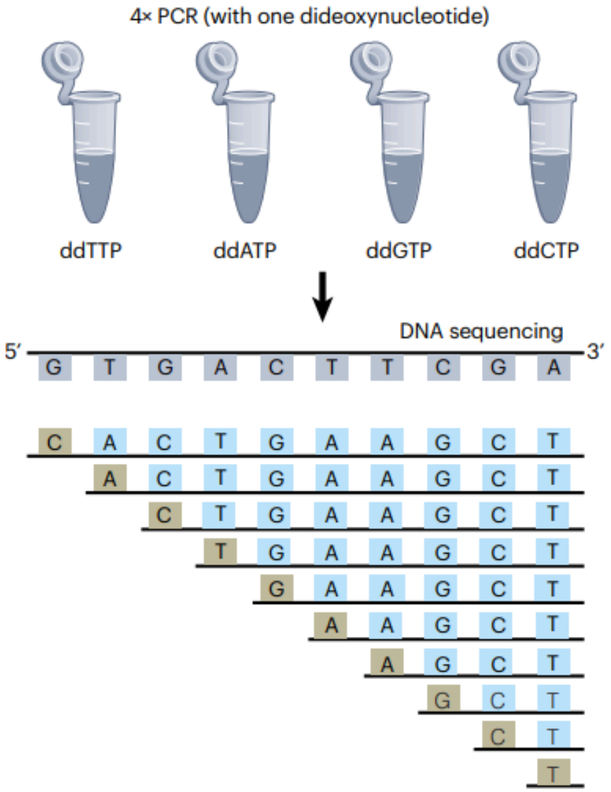
F. SANGER, S. NICKLEN, AND A. R. COULSON

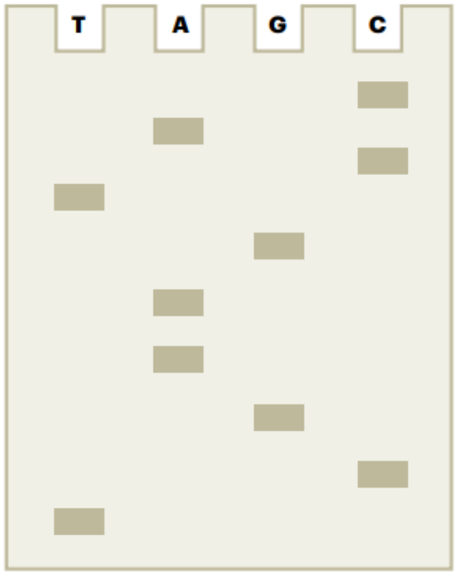
Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 2QH, England

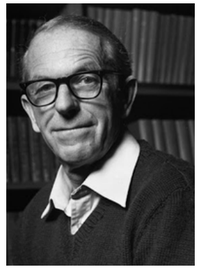
Contributed by F. Sanger, October 3, 1977











DNA sequencing with chain-terminating inhibitors

(DNA polymerase/nucleotide sequences/bacteriophage ϕX174)

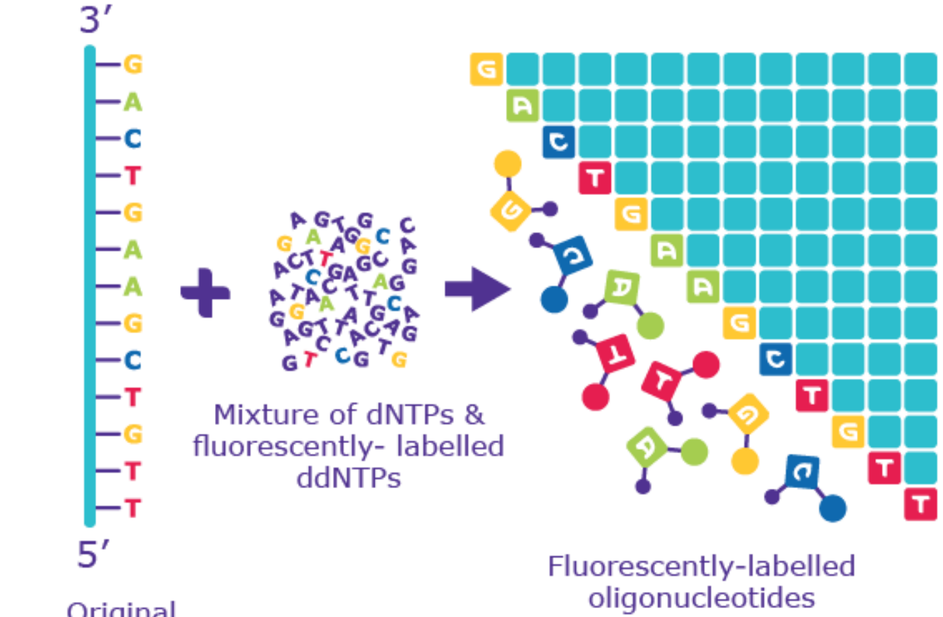
F. SANGER, S. NICKLEN, AND A. R. COULSON

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

Contributed by F. Sanger, October 3, 1977

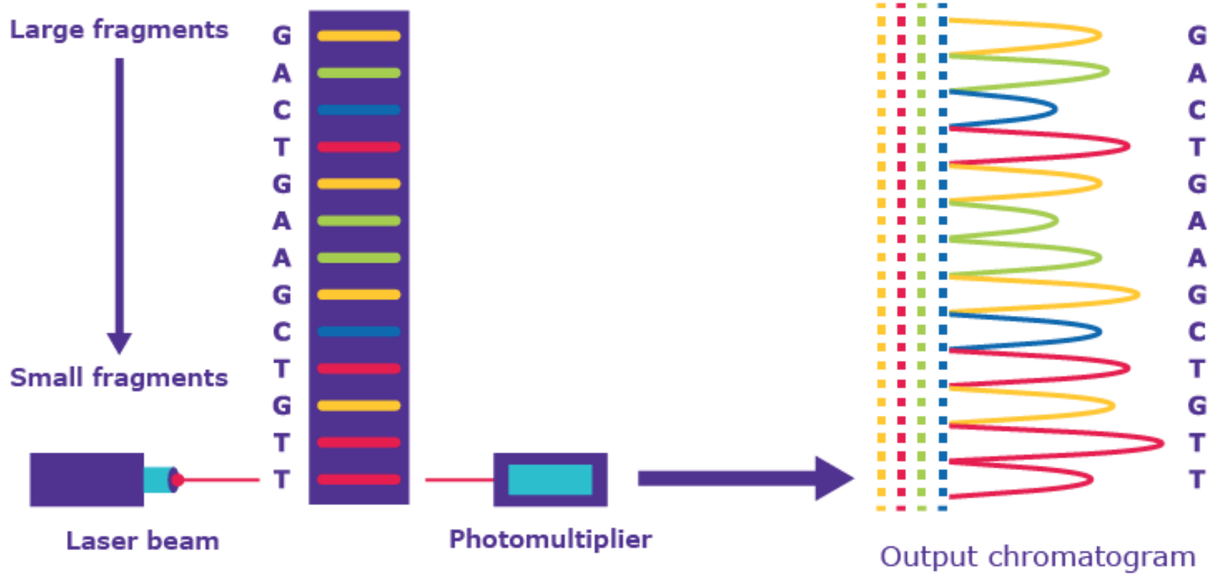
~100 Kb transfer rate per round

1 PCR with fluorescent, chain-terminating ddNTPs



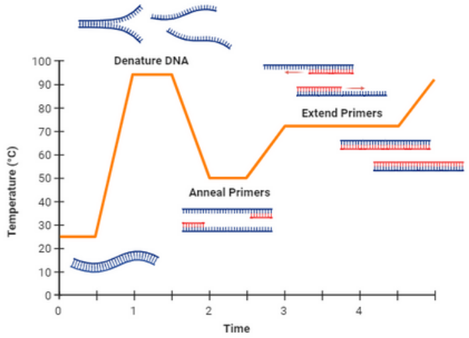
2 Size separation by capillary gel electrophoresis

3 Laser excitation & detection by sequencing machine



Shendure, J., Balasubramanian, S., Church, G. et al. DNA sequencing at 40: past, present and future. Nature 550, 345–353 (2017). https://doi.org/10.1038/nature24286

Kary Mullis inveted polymerase chain reaction (PCR)

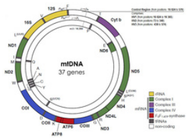


Understanding Our Genetic Inheritance

The U.S. Human Genome Project:

The First Five Years FY 1991-1995

Massively parallel signature sequencing (MPSS) as a tool for in-depth quantitative gene expression profiling in all organisms



1981

1982

NCBI

1983

1987

Applied Biosystems, ABI370. 1,000 bases per day

1990

Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd

1995

1996

Life with 6000 Genes

A. Goffeau,* B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, S. G. Oliver

Saccharomyces cerevisiae

1995

2000

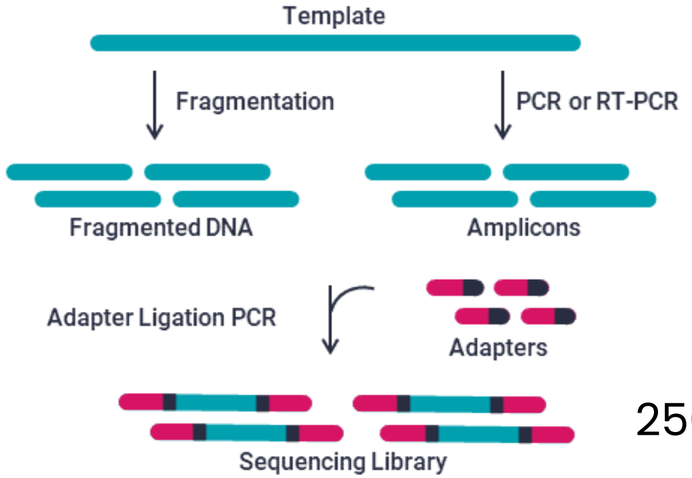
Arabidopsis thaliana

2002

2ª Geração de Sequenciamento

STEP 1:
Extraction

STEP 2:
Library Prep



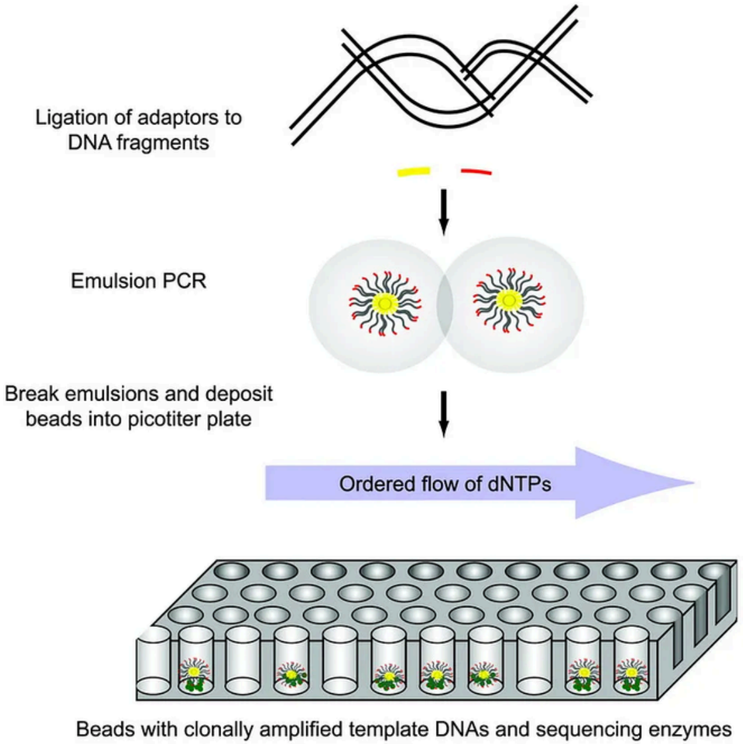
250– 800 bp

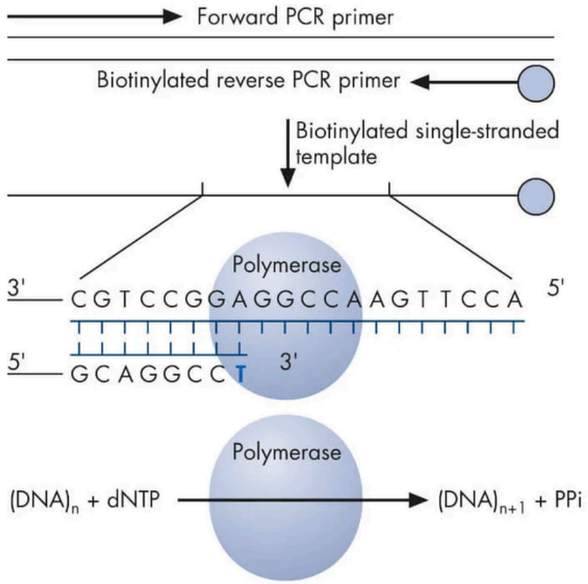
STEP 3:
Sequencing

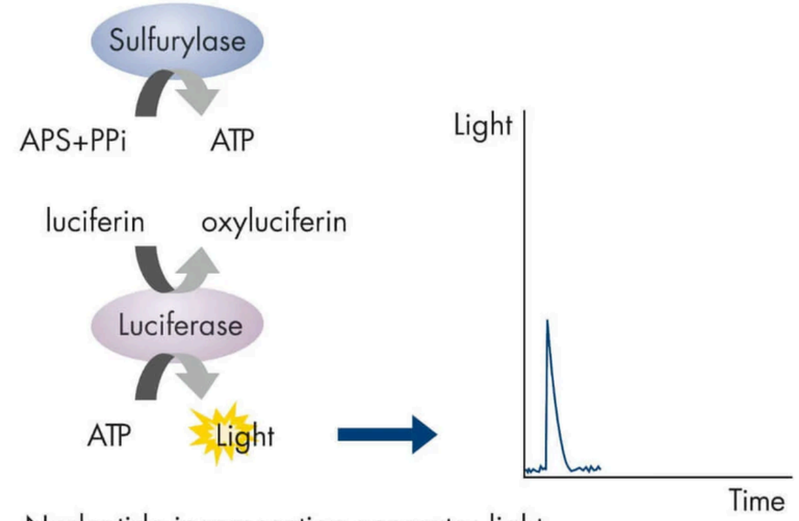
STEP 4:
Analysis

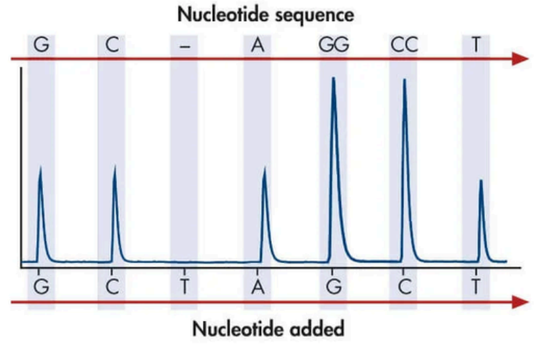
Owen-Hughes T, Engeholm M. Pyrosequencing positions nucleosomes precisely.
Genome Biol. 2007;8(6):217. doi: 10.1186/gb-2007-8-6-217. PMID: 17601355; PMCID: PMC2394743.

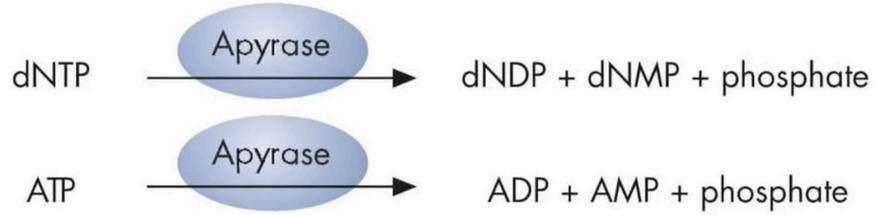
Principle of Pyrosequencing







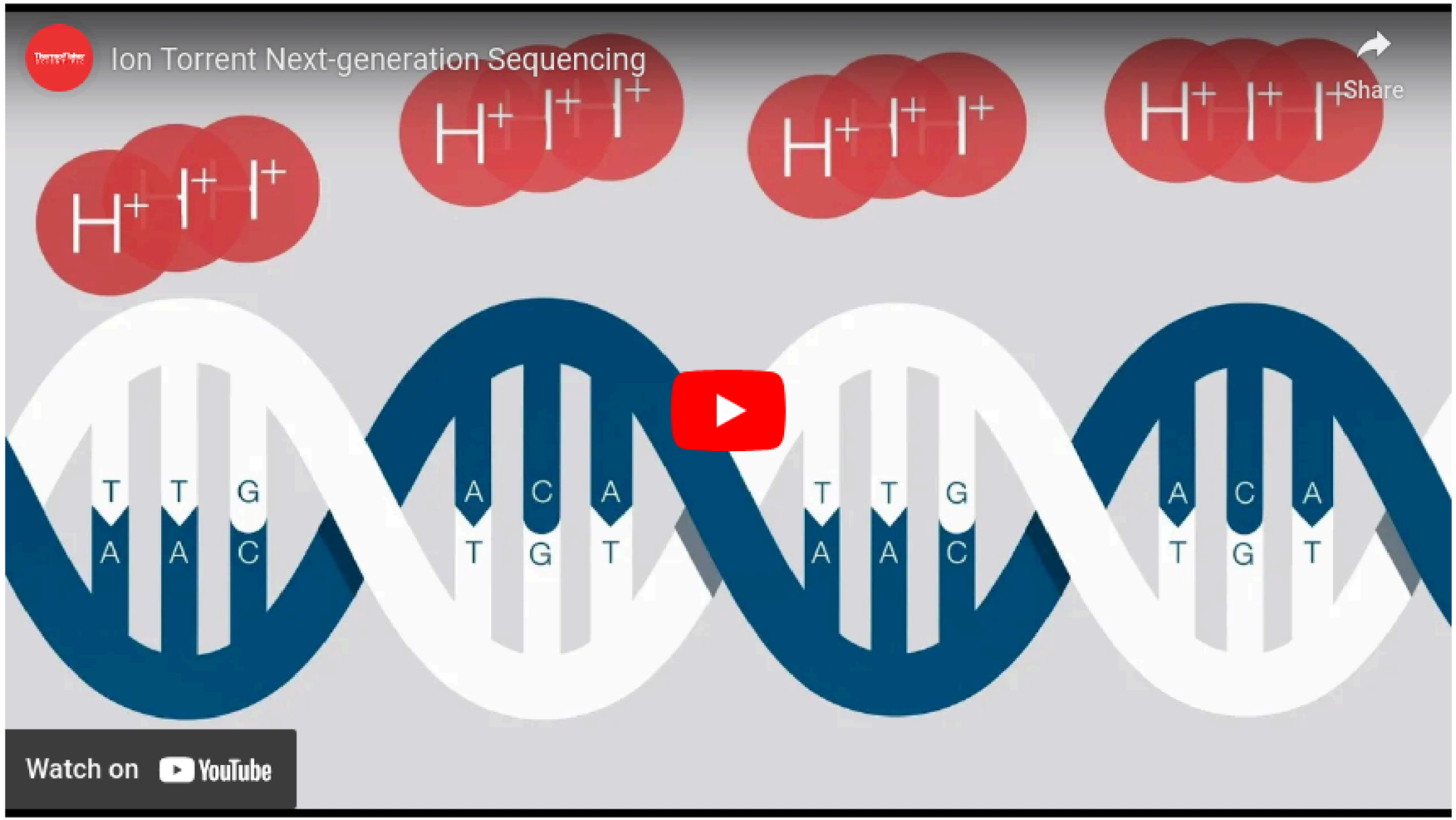




QIAGEN

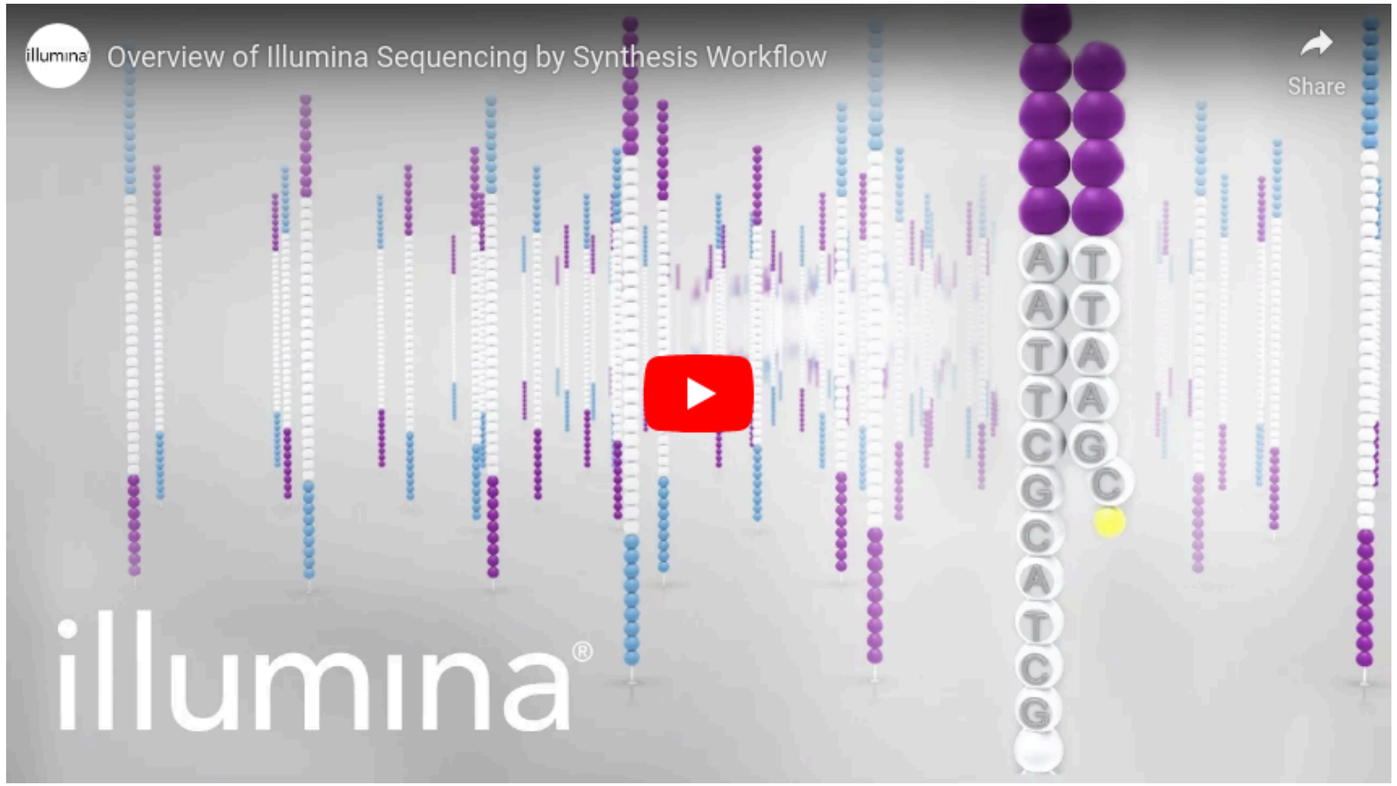
454 Roche GS FLX System (2004)
400-600 million base pairs per run

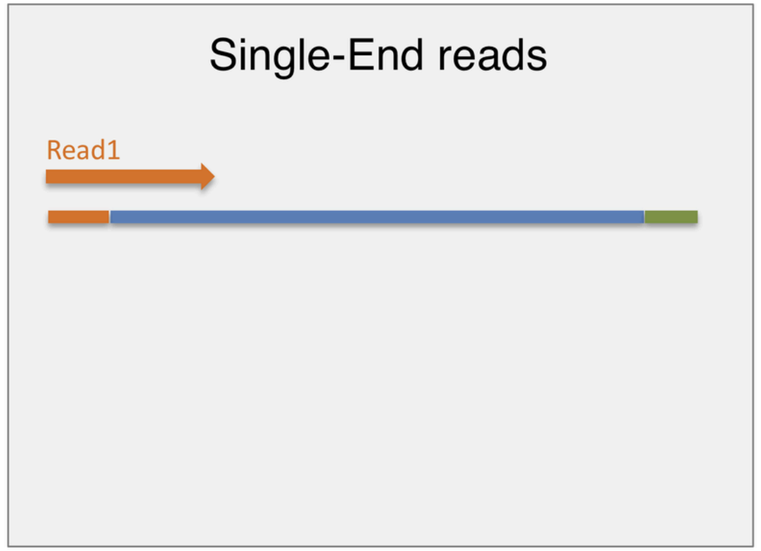
SBS - Sequencing by detection of hydrogen ions

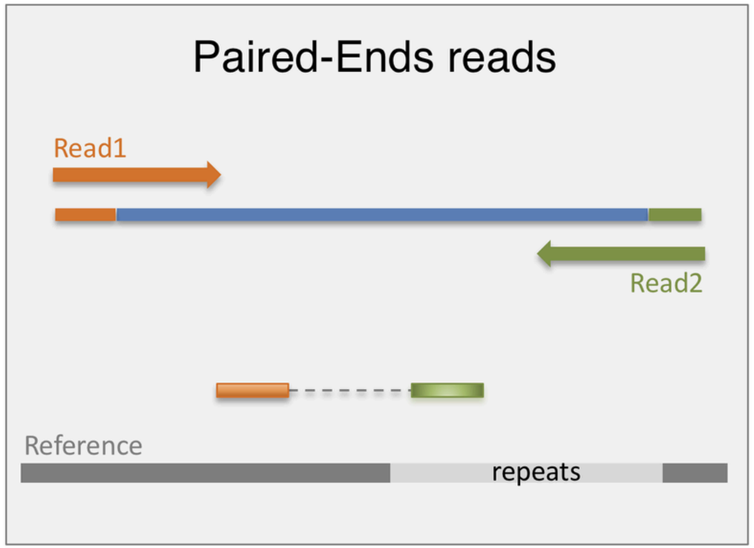


Sequencing, by synthesis (SBS)

SBS - Sequencing by reversible termination

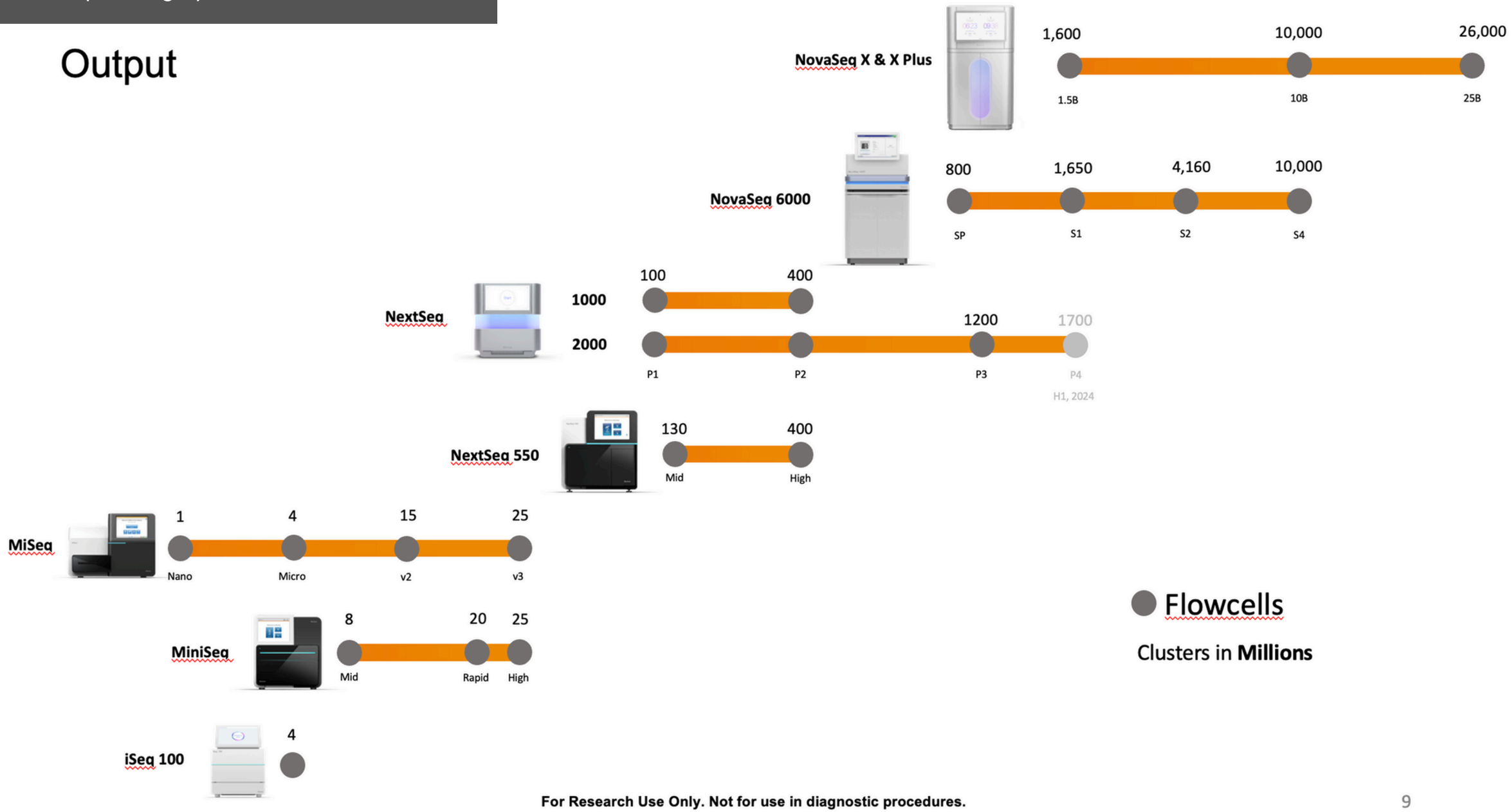




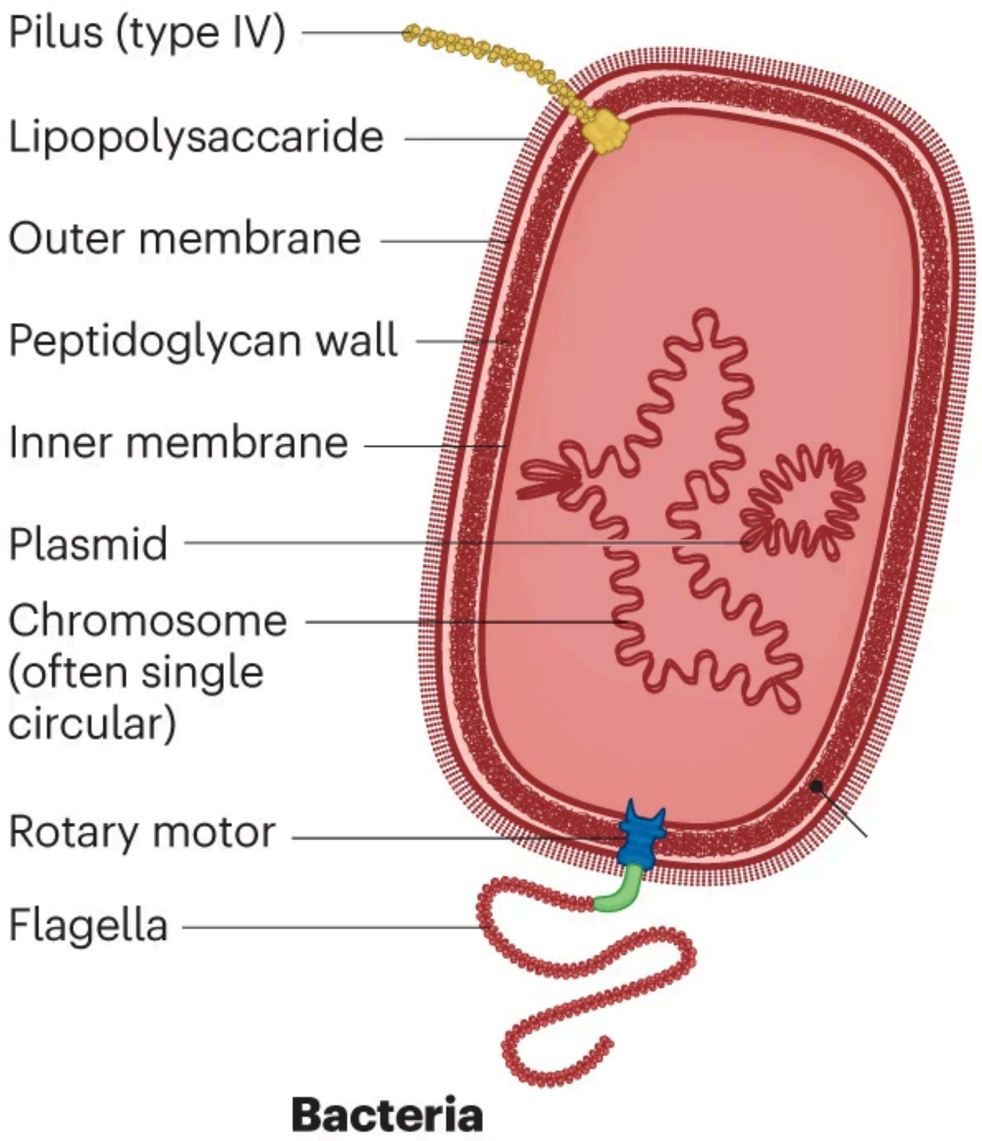


Sequencing, by synthesis (SBS)

Output



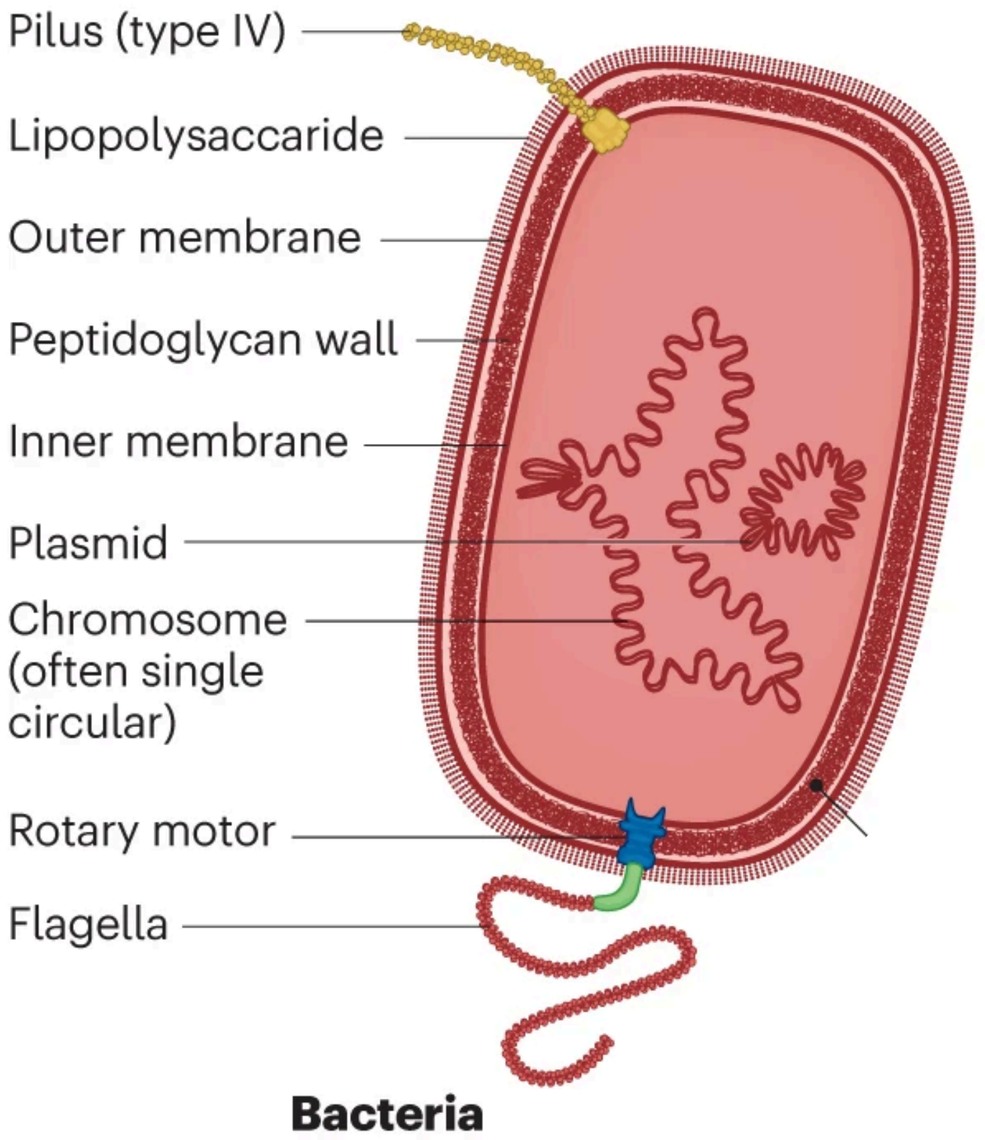
Genoma dos procariotos



ESTRUTURA DO GENOMA

PROCARIOTOS

- Separação física clara entre o DNA e o citoplasma
- Replicação, segregação e duplicação celular estão fortemente interligadas
- Proporcionalidade entre o tamanho do genoma e o número de proteínas codificadas
- Os plasmídeos carregam genes para sua propagação, manutenção na célula, adaptação



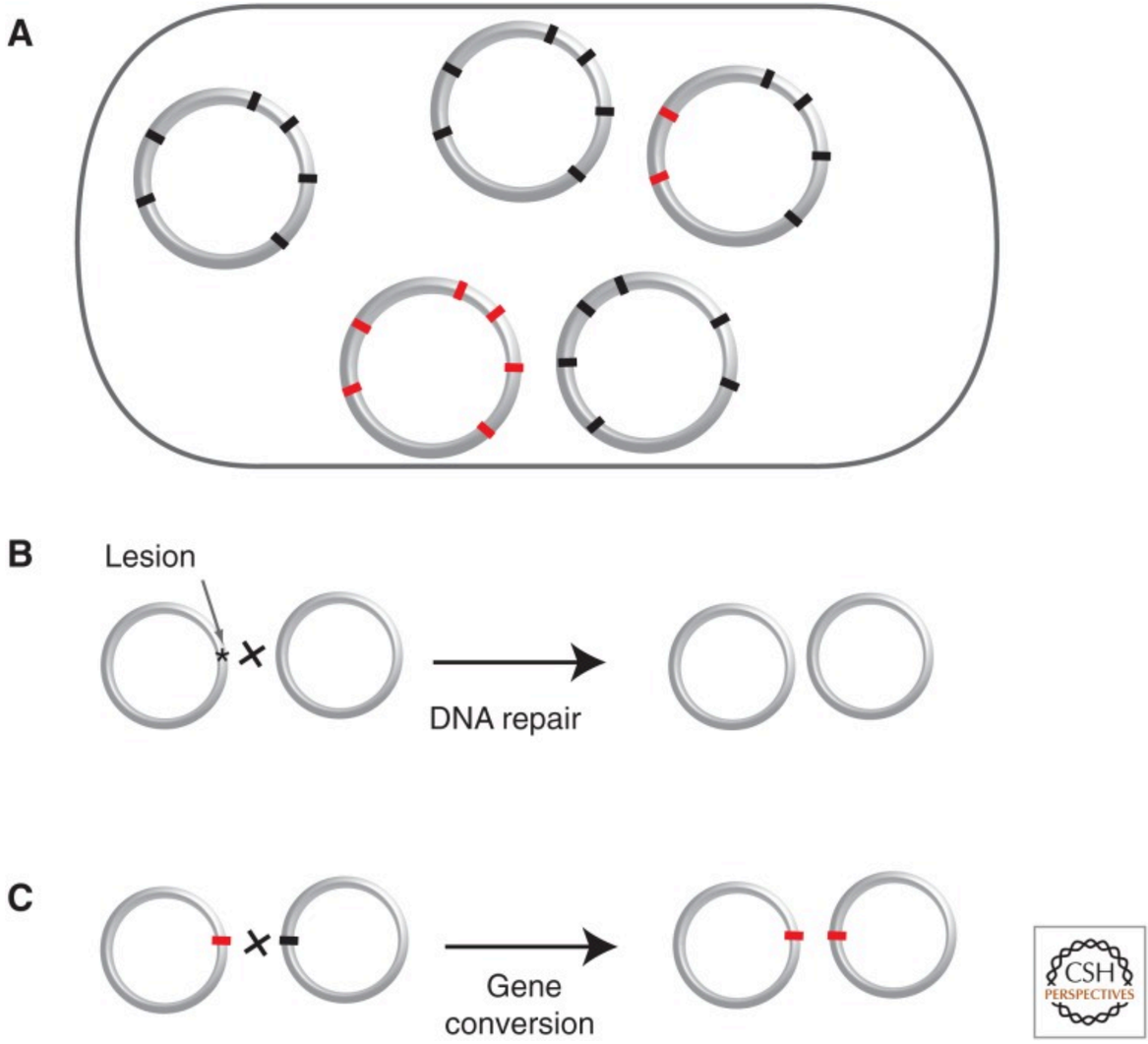
https://www.nature.com/articles/s41576-023-00688-5/figures/1

https://cshperspectives.cshlp.org/content/8/1/a018168.long

ESTRUTURA DO GENOMA

PROCARIOTOS

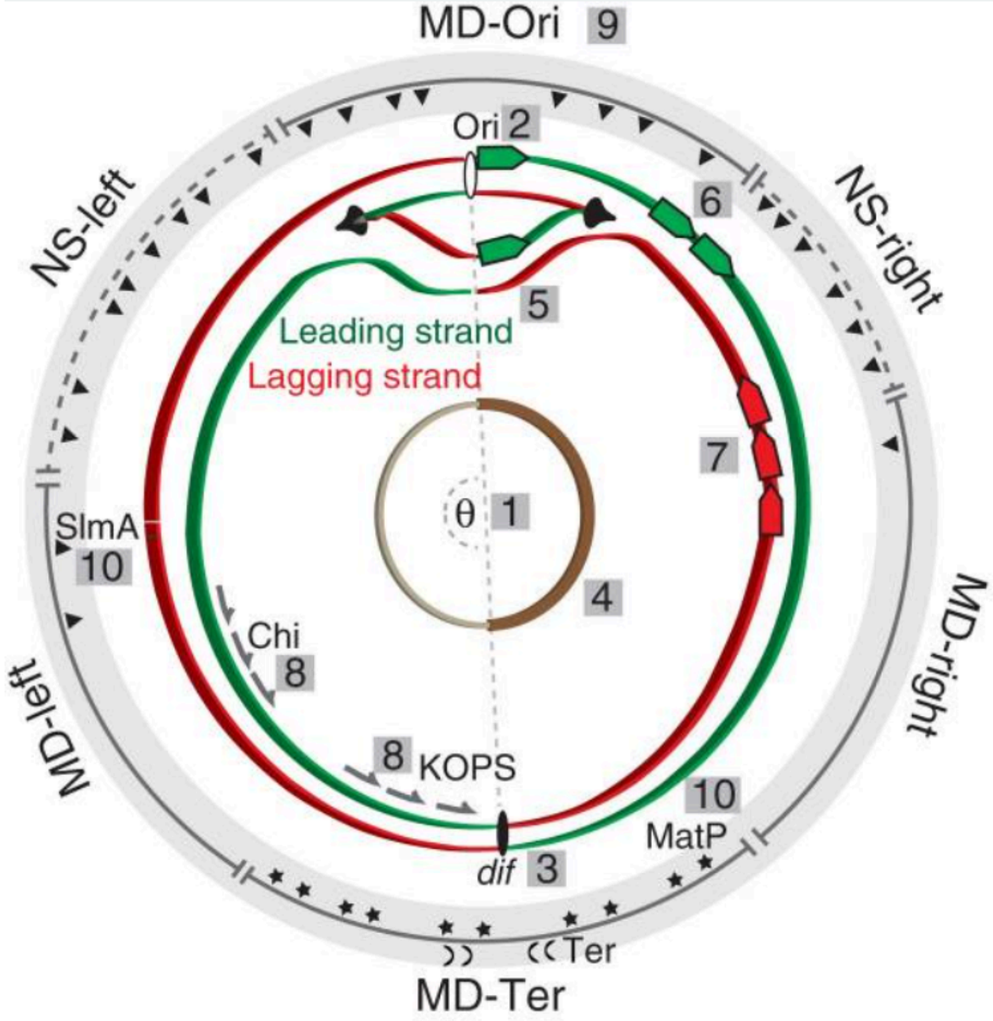
- A poliploidia pode estar relacionada com a expressão gênica, a reparação de DNA ou a eficiência da seleção natural



https://cshperspectives.cshlp.org/content/8/1/a018168.long

ESTRUTURA DO GENOMA

PROCARIOTOS



1 Chromosomes are **symmetric** ($\theta \sim 180^{o}$)

2 Replication starts at the origin (**Ori**)

3 Replication terminates close to the ***dif*** site

4 **GC skews** are higher in the leading strand

5 Highly expressed genes cluster at Ori for replication-associated **gene dosage effects**

6 **Gene strand bias** results in more genes co-oriented with replication fork progression

7 Functionally neighbor genes are cotranscribed in **operons**

8 **Leading strand overrepresents DNA motifs** implicated in repair (Chi) or in segregation (KOPS)

9 Chromosomes are organized in **structured macrodomains** (MD) **and unstructured flexible regions** (NS)

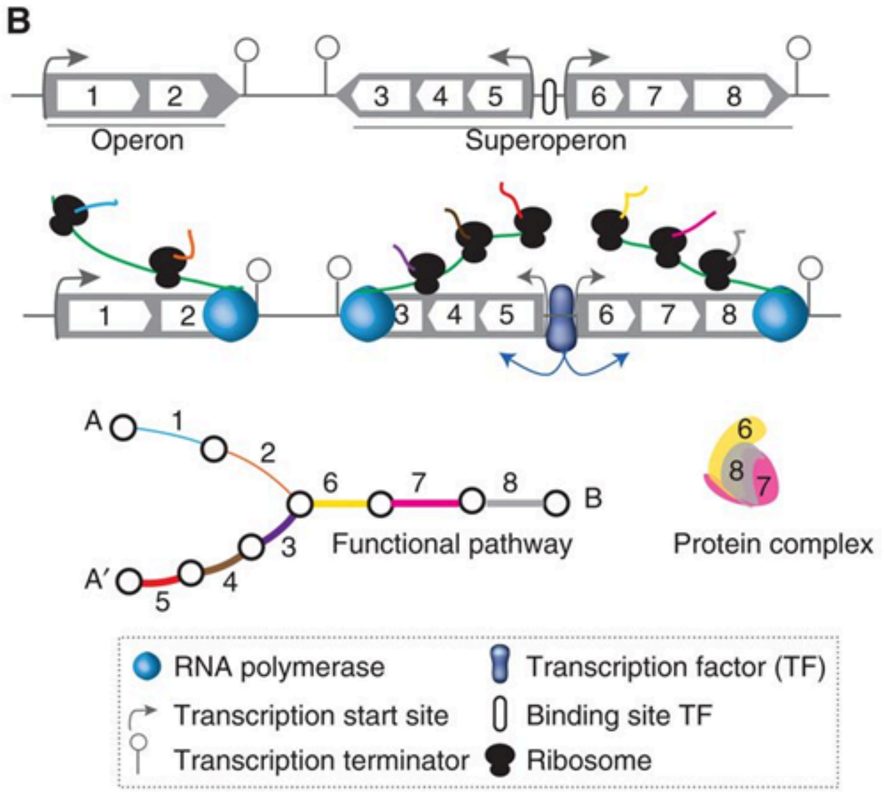
10 Proteins with **specific DNA-binding** properties drive **nucleoid dynamics**

https://cshperspectives.cshlp.org/content/8/1/a018168.long

ESTRUTURA DO GENOMA

PROCARIOTOS

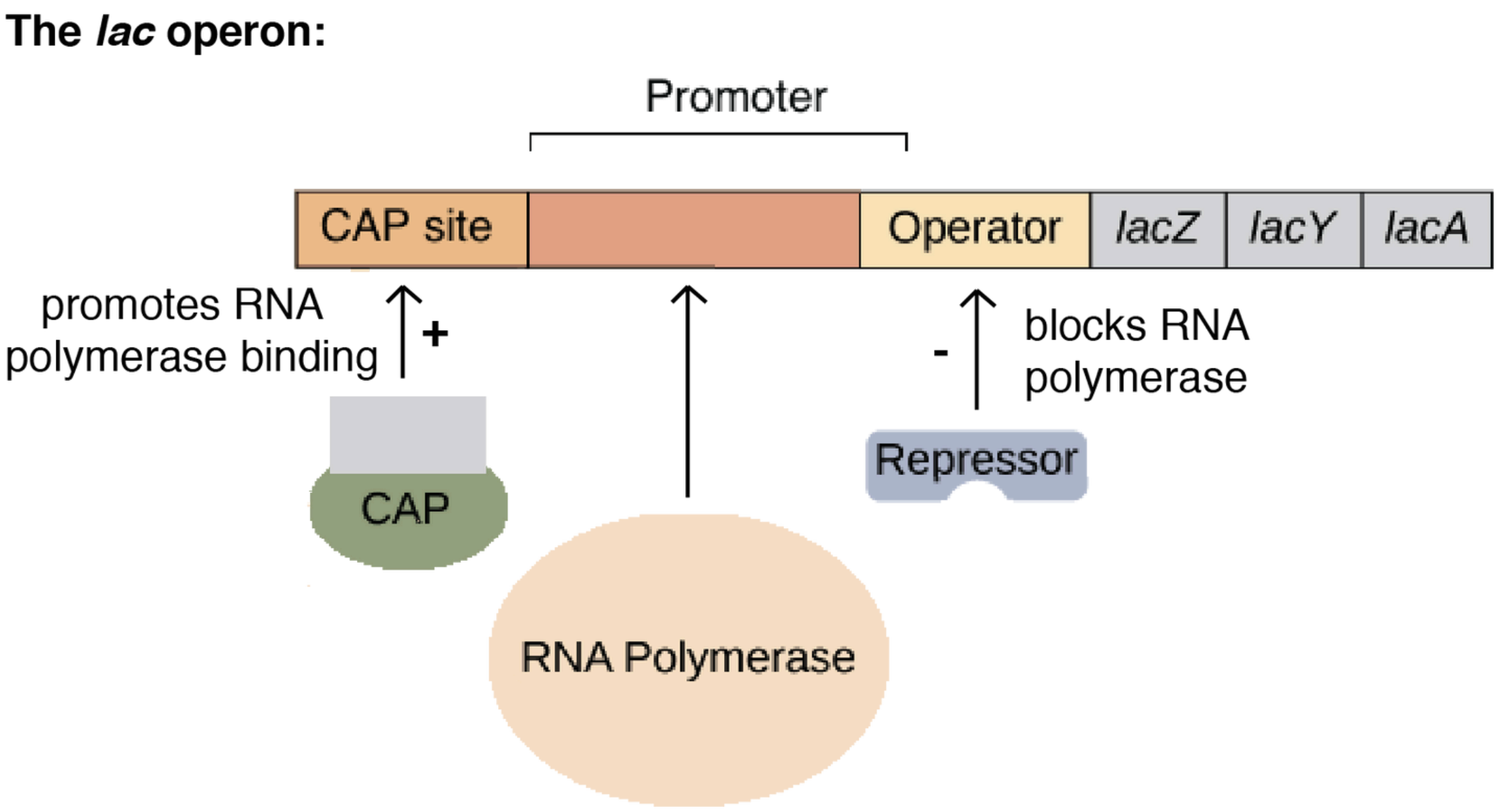
- Operons são conjuntos de genes sob o controle de um único sítio de início de transcrição
- A maioria dos genes em procariotos é expressa na forma de operons
- Pares de genes contíguos em operons são altamente conservados
- Os genes em operons frequentemente codificam proteínas que interagem fisicamente ou vizinhos funcionais



https://cshperspectives.cshlp.org/content/8/1/a018168.long

ESTRUTURA DO GENOMA

PROCARIOTOS

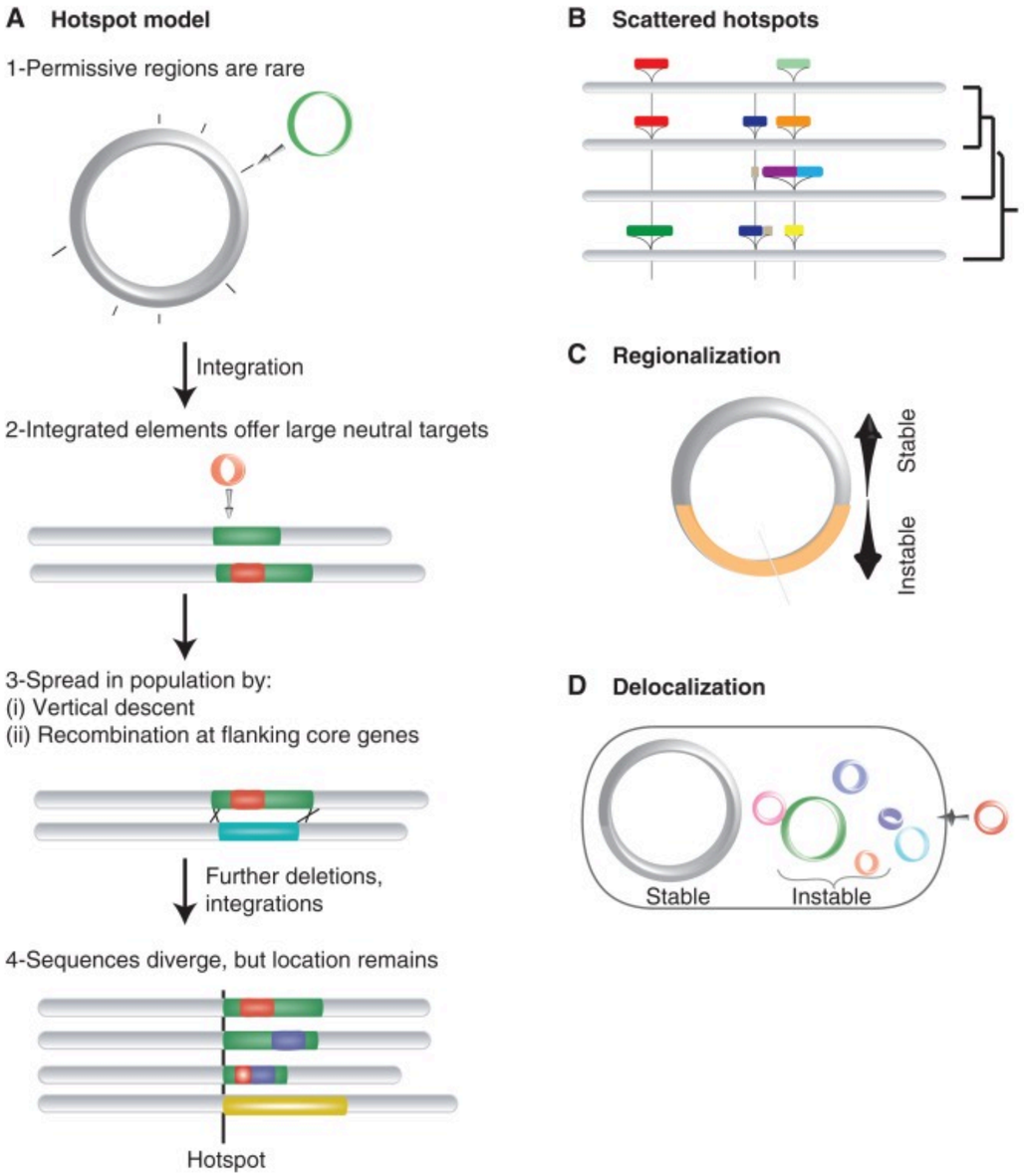


https://www.sciencedirect.com/science/article/pii/S0022283661800727?via%3Dihub

ESTRUTURA DO GENOMA

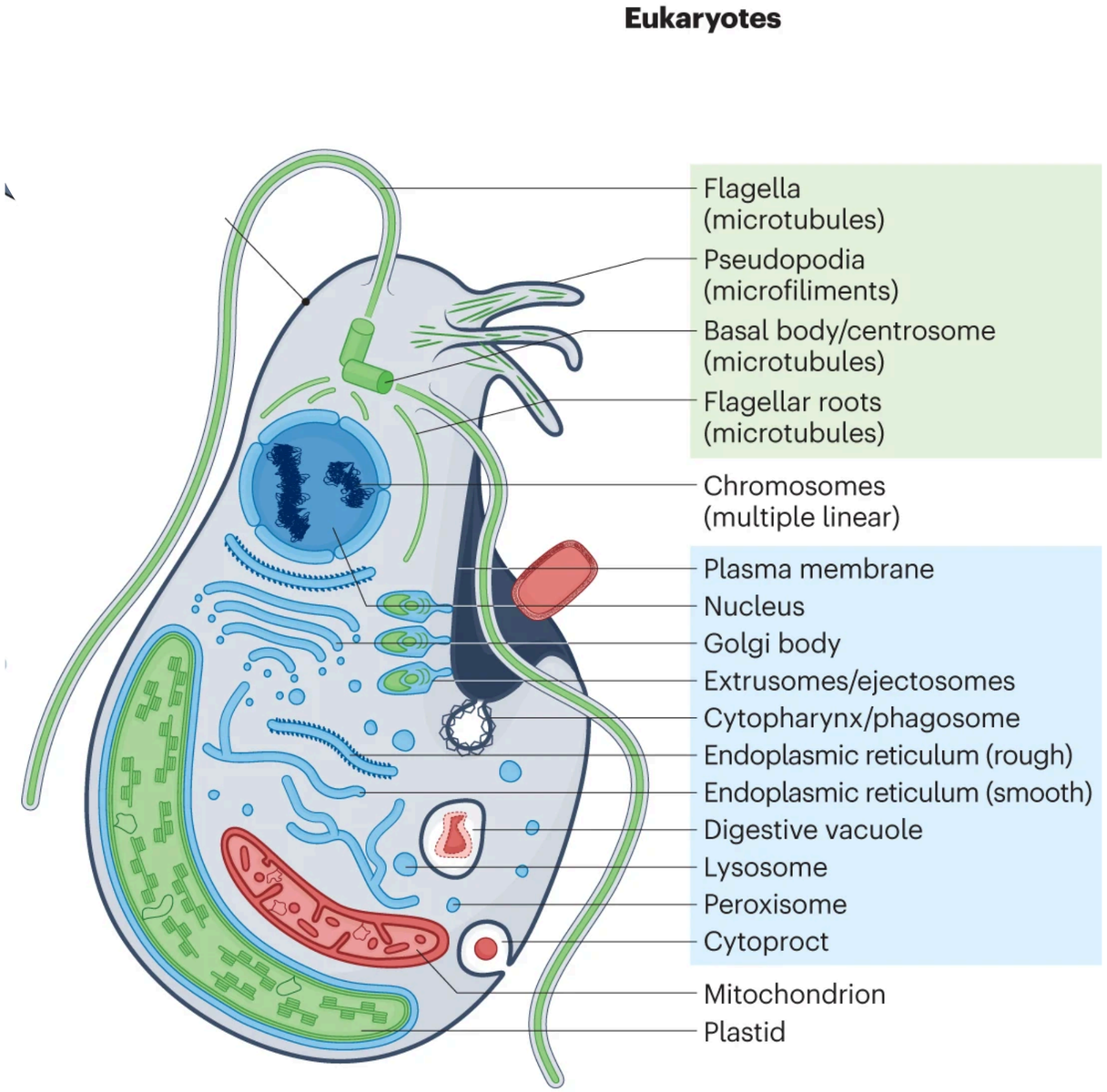
PROCARIOTOS

- As características organizacionais são fortemente afetadas por rearranjos no genoma, -> cromossomos assimétricos, romper operons e desorganizar domínios cromossômicos
- Aquisição de novos genes (THG)



https://cshperspectives.cshlp.org/content/8/1/a018168.long

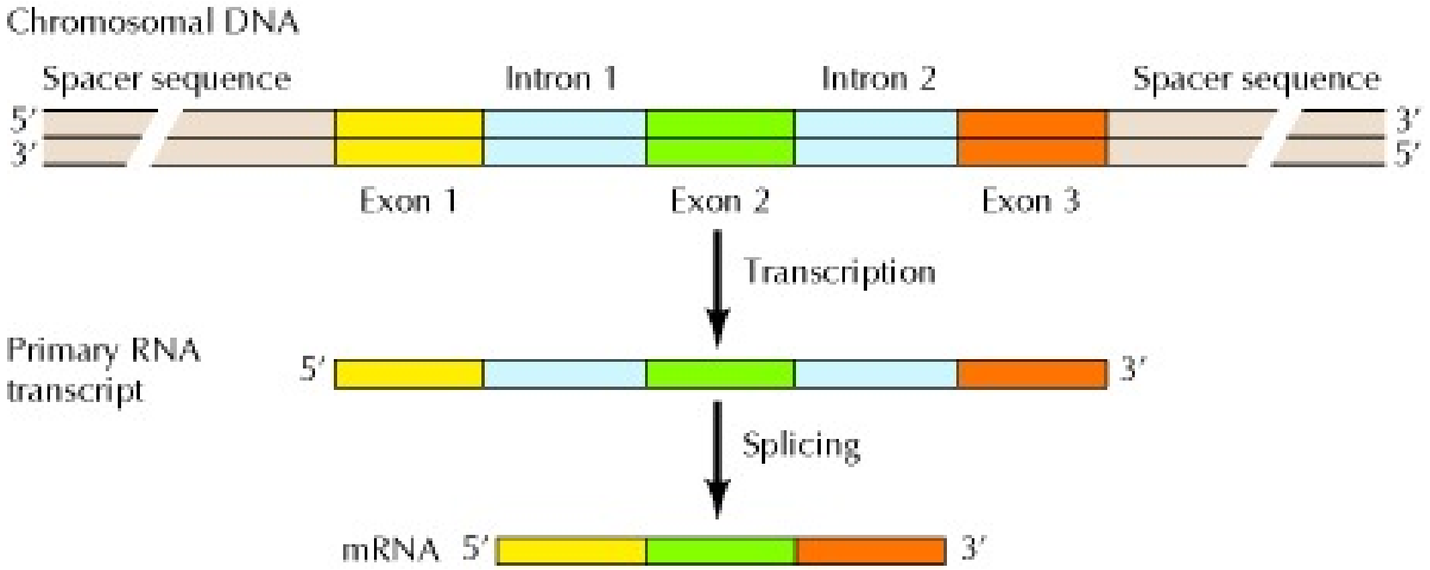
Genoma dos eucariotos

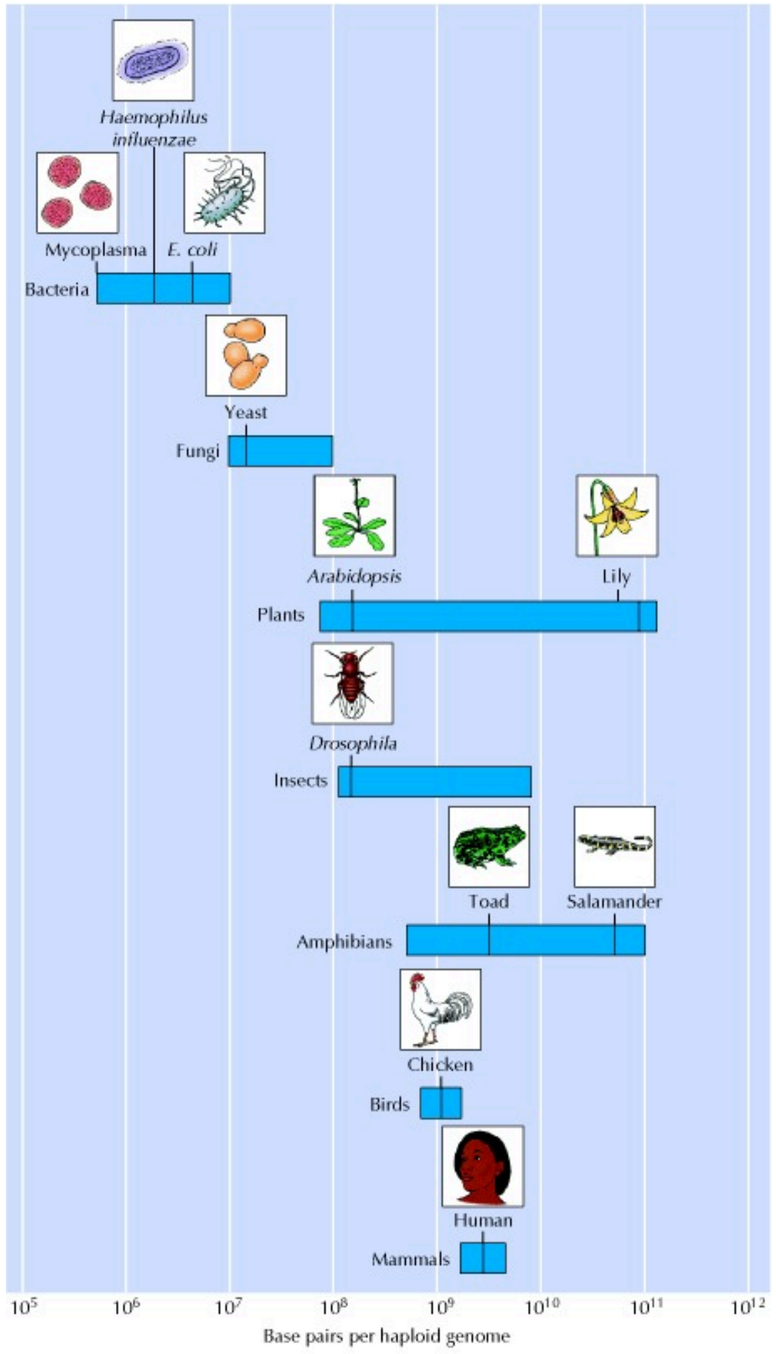


COMPLEXIDADE DO GENOMA

EUCARIOTOS

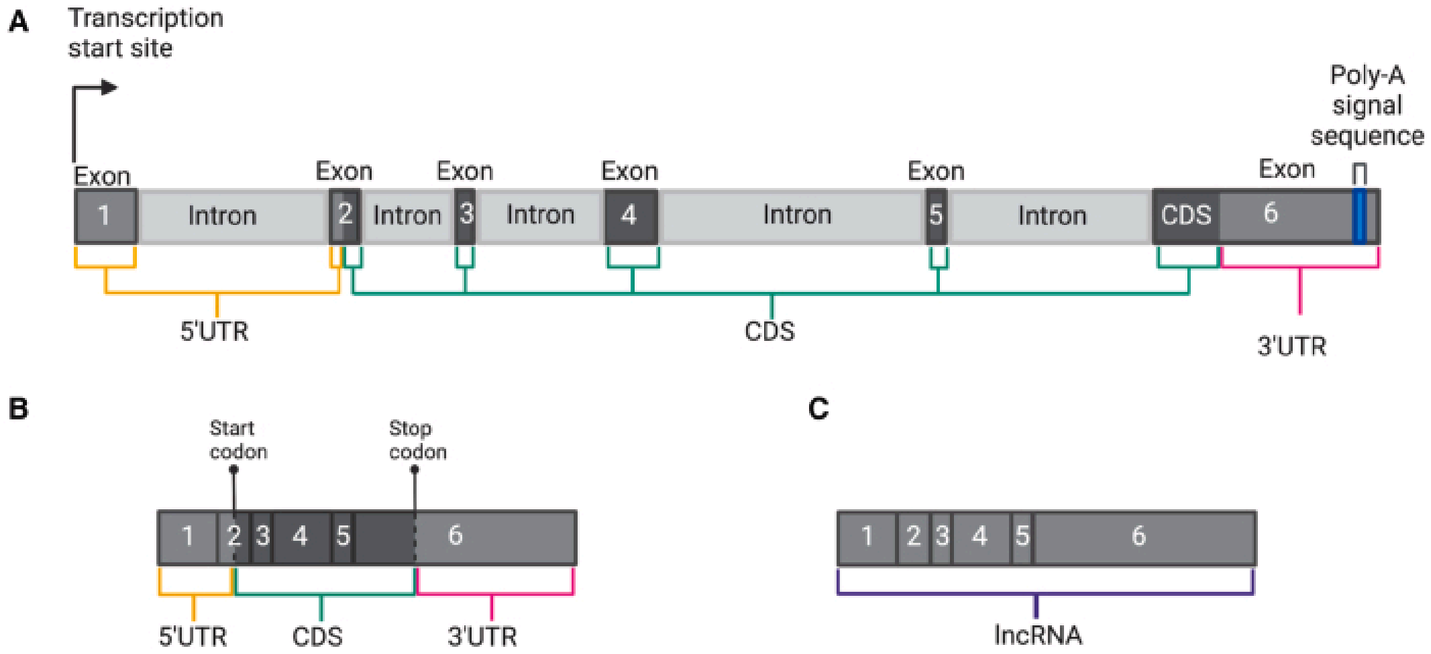
Grande parte da complexidade dos genomas eucarióticos resulta da abundância de sequências não codificantes



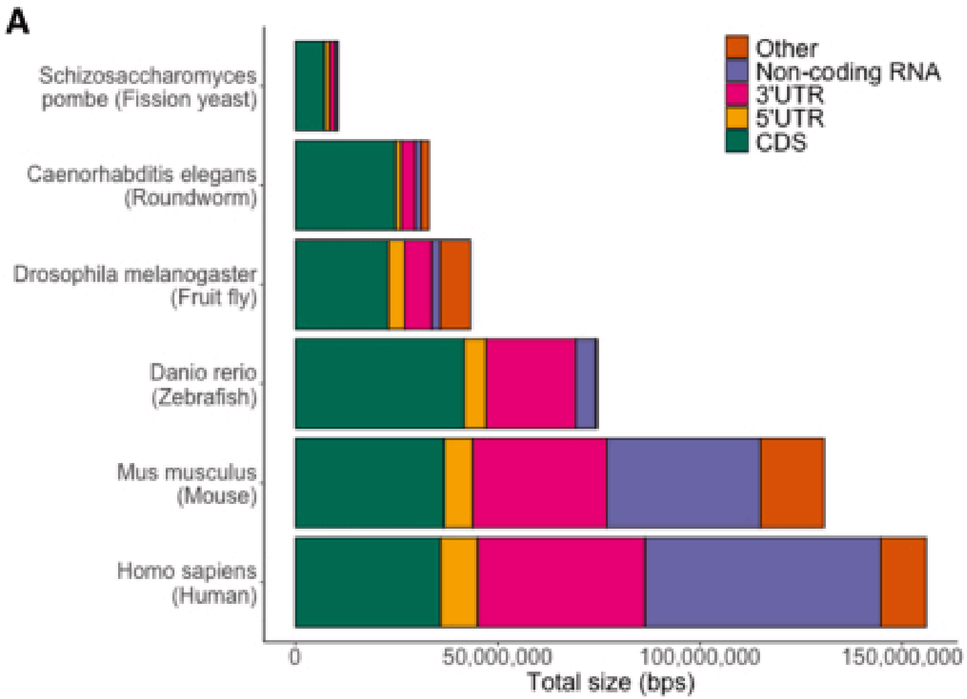


COMPLEXIDADE DO GENOMA

NEM TODO ÉXON CODIFICA PROTEÍNA



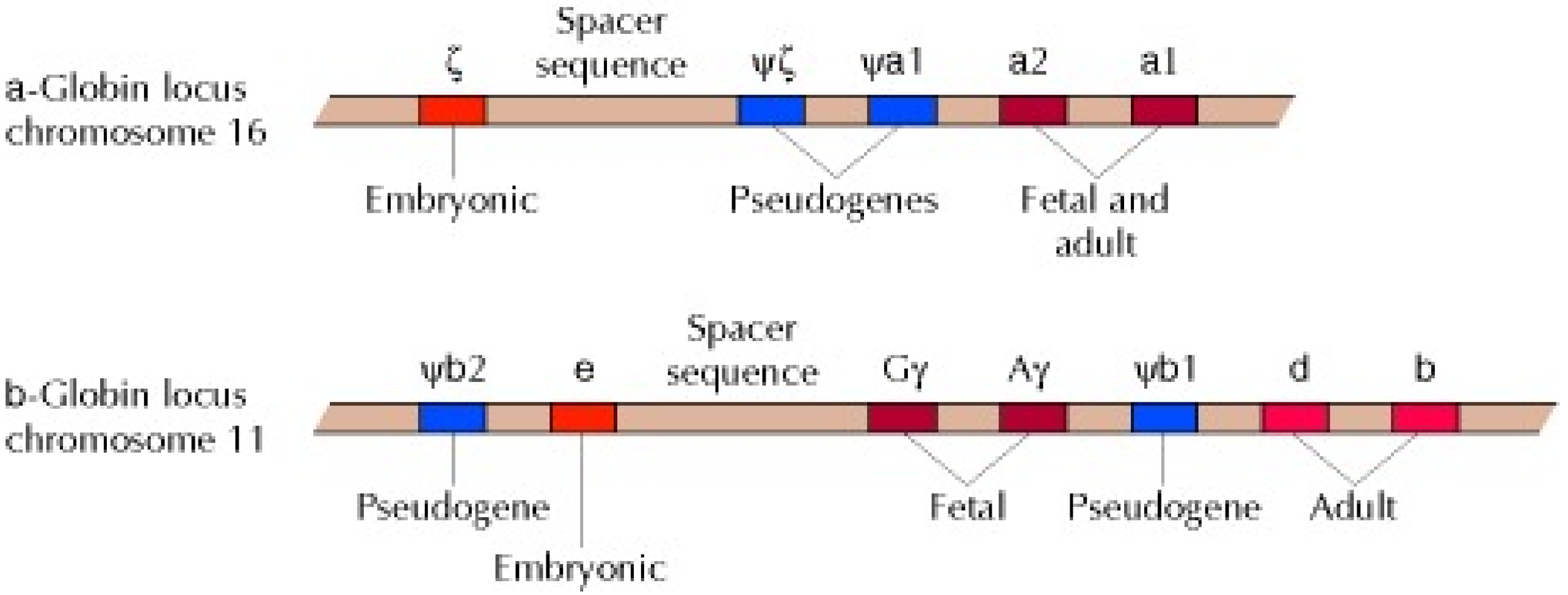
Proporção de sequências exônicas e representação no sequenciamento do exoma completo



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10112331/

COMPLEXIDADE DO GENOMA

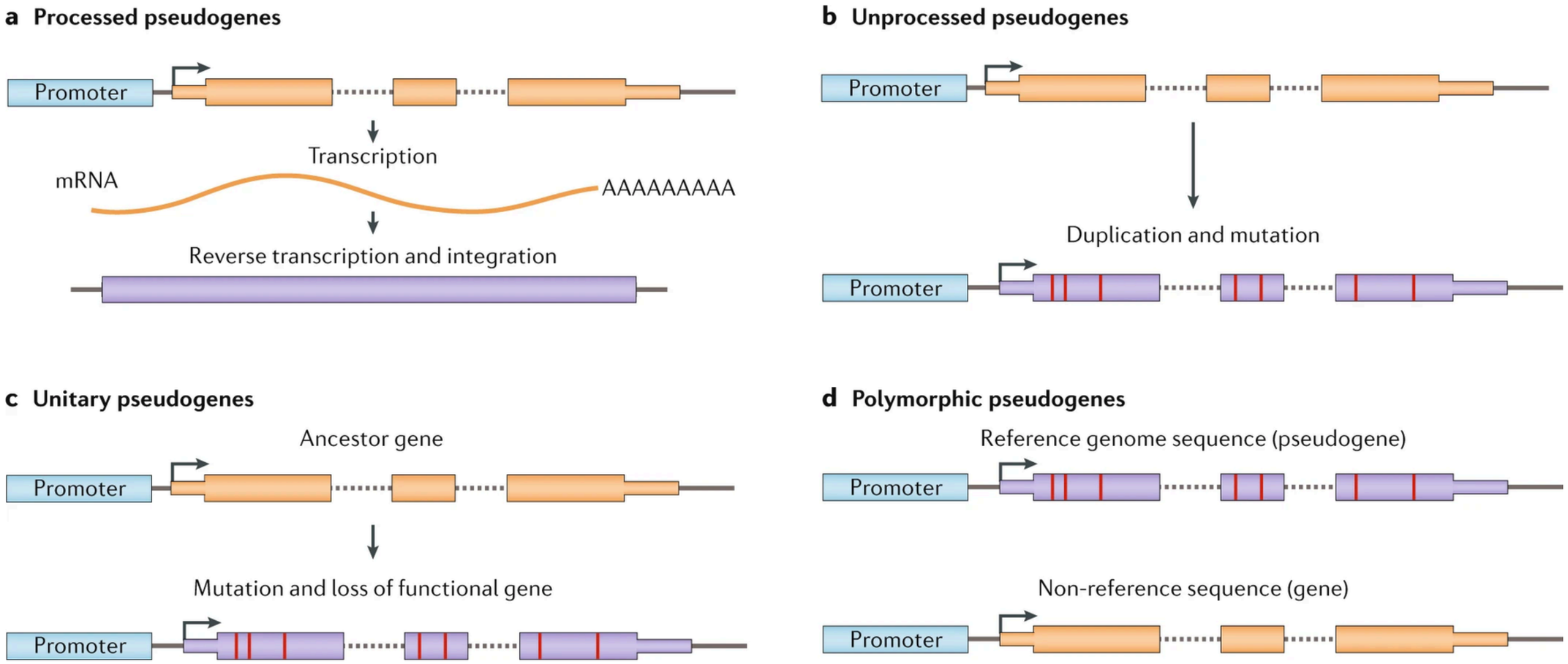
FAMÍLIA DE GENES



https://www.ncbi.nlm.nih.gov/books/NBK9846/

COMPLEXIDADE DO GENOMA

PSEUDOGENES

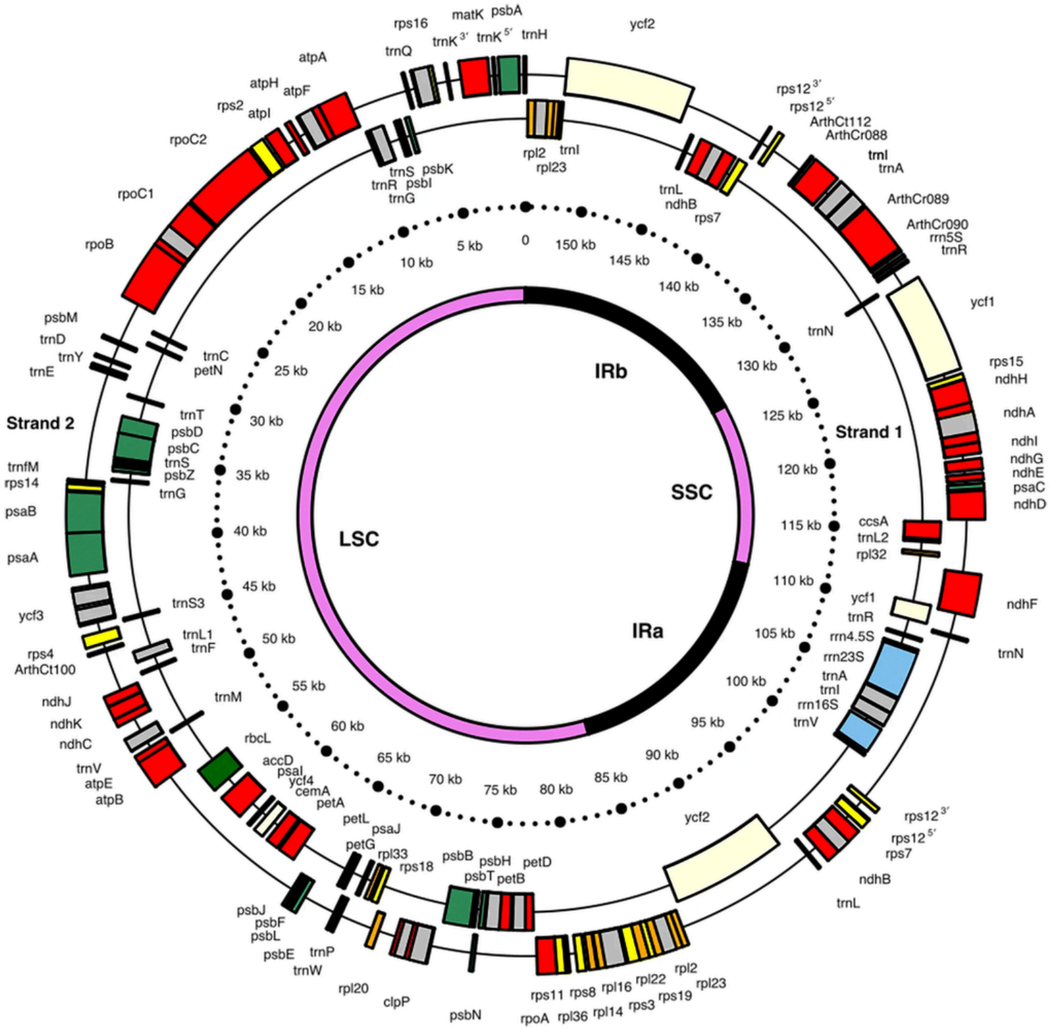


https://www.nature.com/articles/s41576-019-0196-1

COMPLEXIDADE DO GENOMA

PLASTOMA

- Inicialmente, os estudos de sequenciamento se limitavam a pequenos trechos de DNA (barcoding)
- Estudos filogenéticos
- gene COI (mitocondrial)



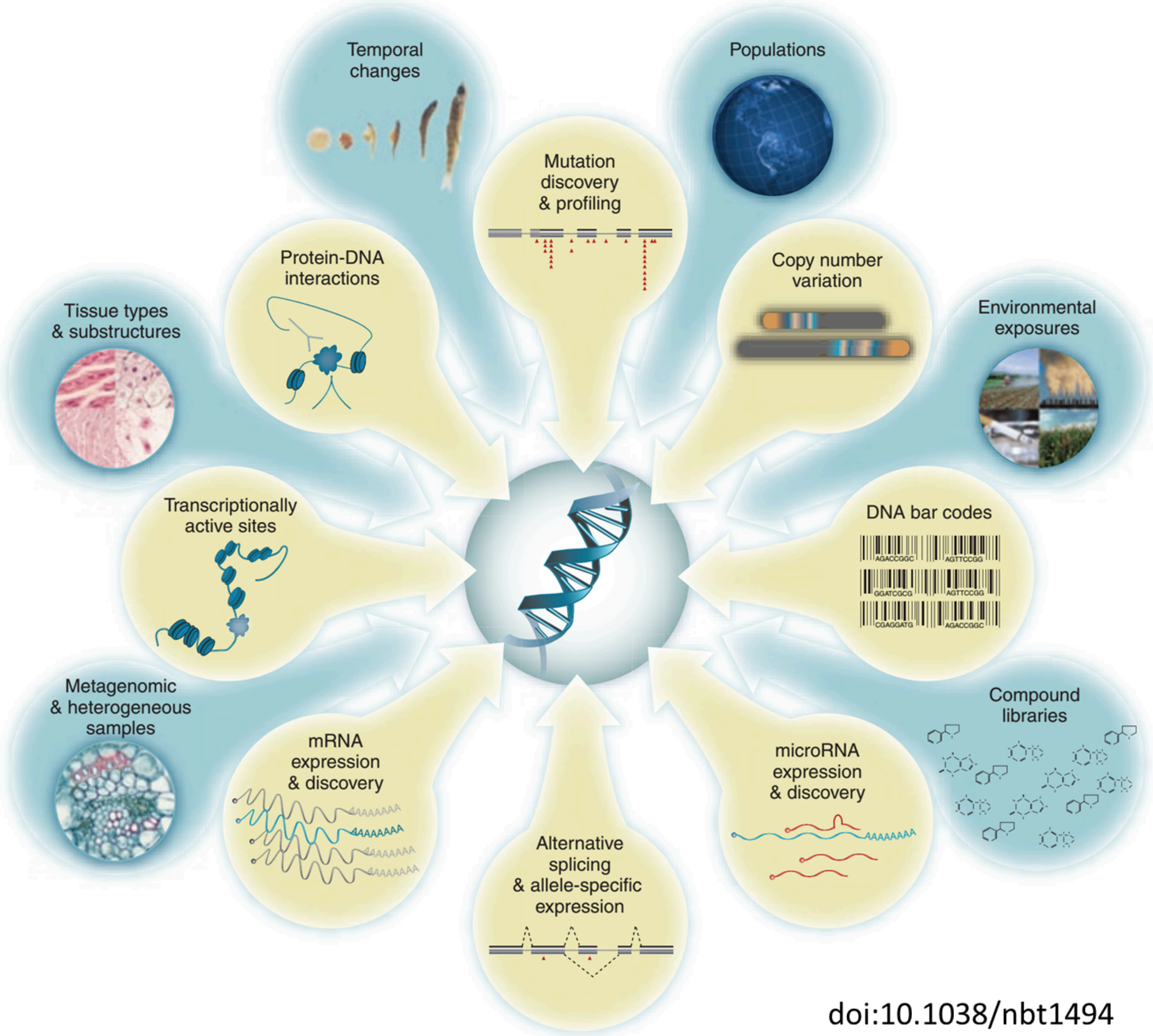
COMPLEXIDADE DO GENOMA

REPETITOMA

- “Junk DNA” -> não codificadoras de proteínas
- Repetições em tandem -> microssatélites
- Transposons - “genes saltadores”
- Regulação genética, organização cromossômica e evolução

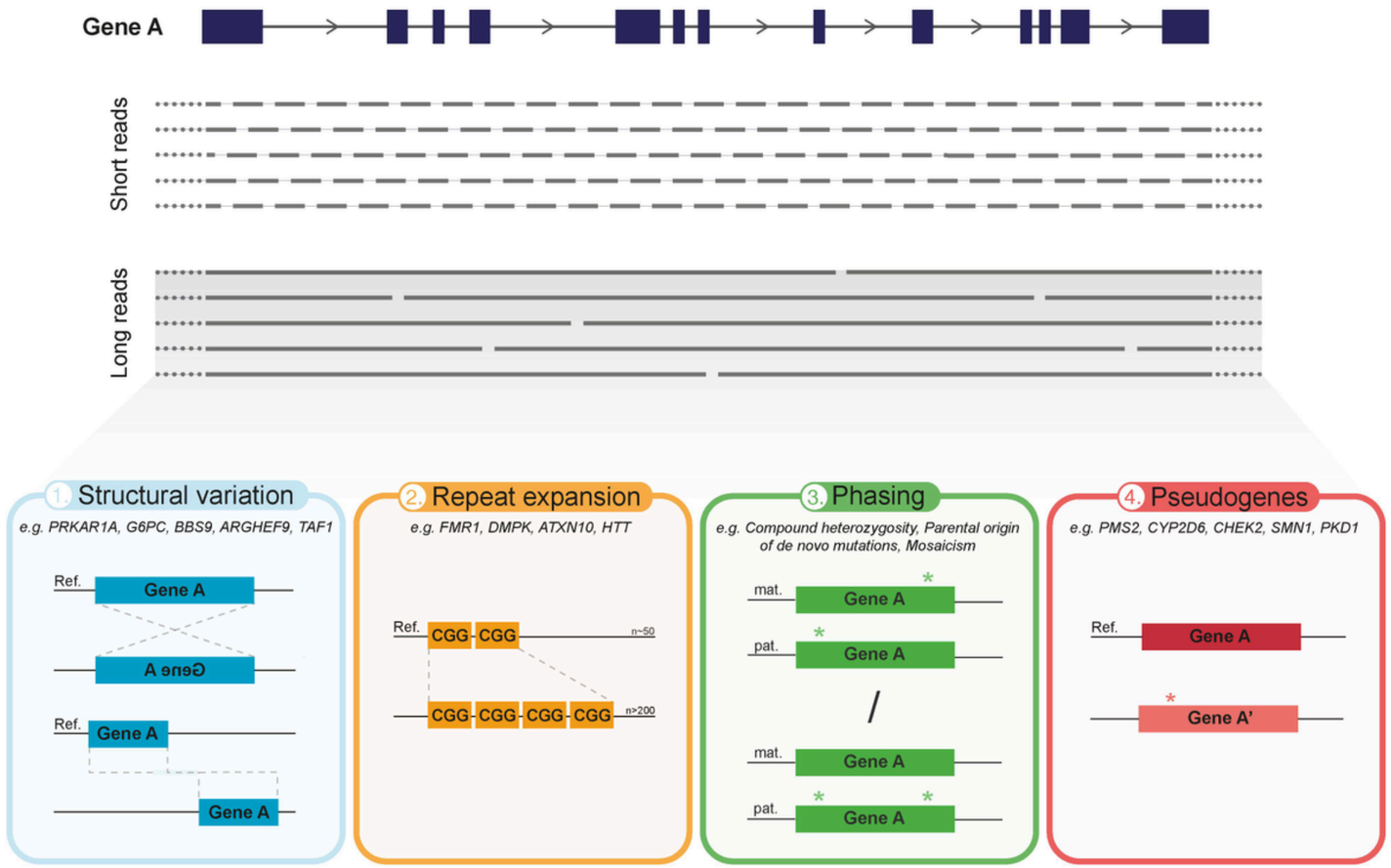
https://www.nature.com/articles/443521a

Advantages and Disadvantages of using NGS Short-Reads



doi:10.1038/nbt1494

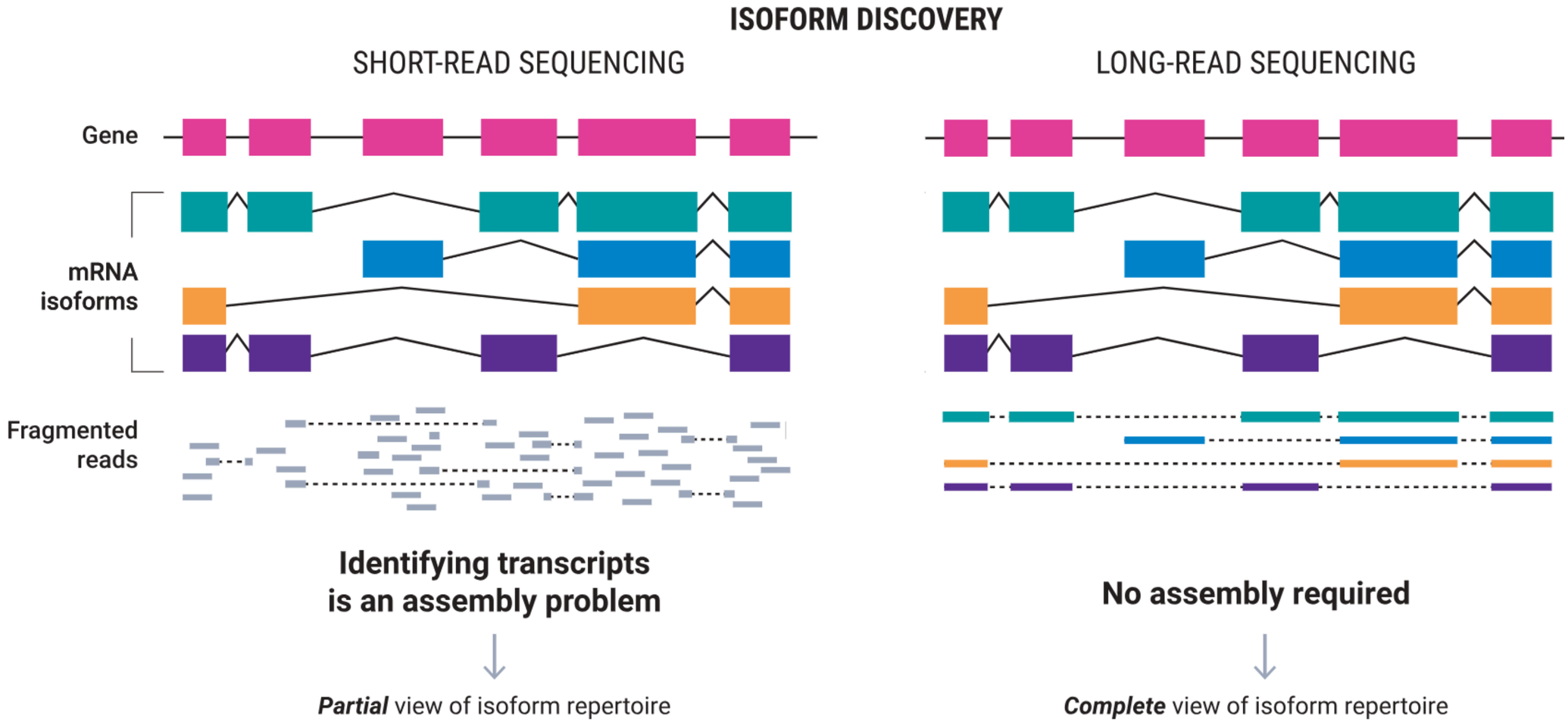
Advantages and Disadvantages of using NGS Short-Reads

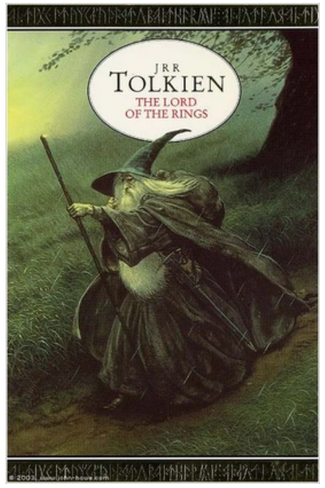


Nearly all of the aforementioned platforms require template amplification. However, the downsides of amplification include **copying errors**, **sequence-dependent biases** and information loss (for example, **methylation)**

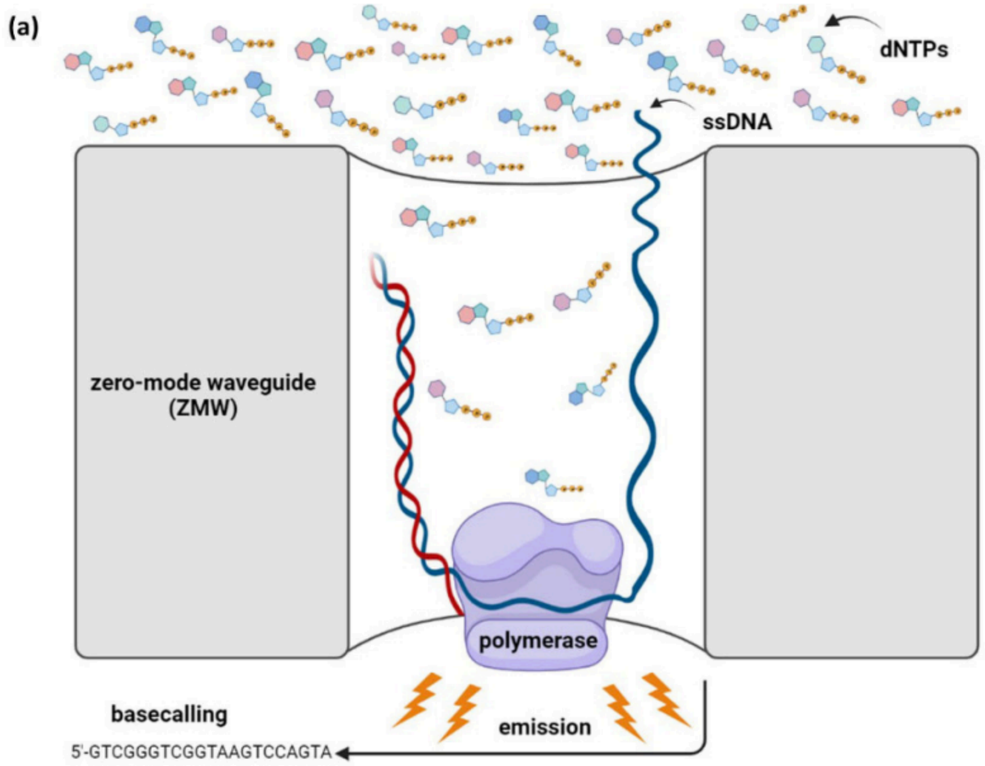
3ª Geração de Sequenciamento

Single molecule real time (SMRT) - 1kb-20kb





Pacific Biosciences Single Molecule Real Time (SMRT) Sequencing



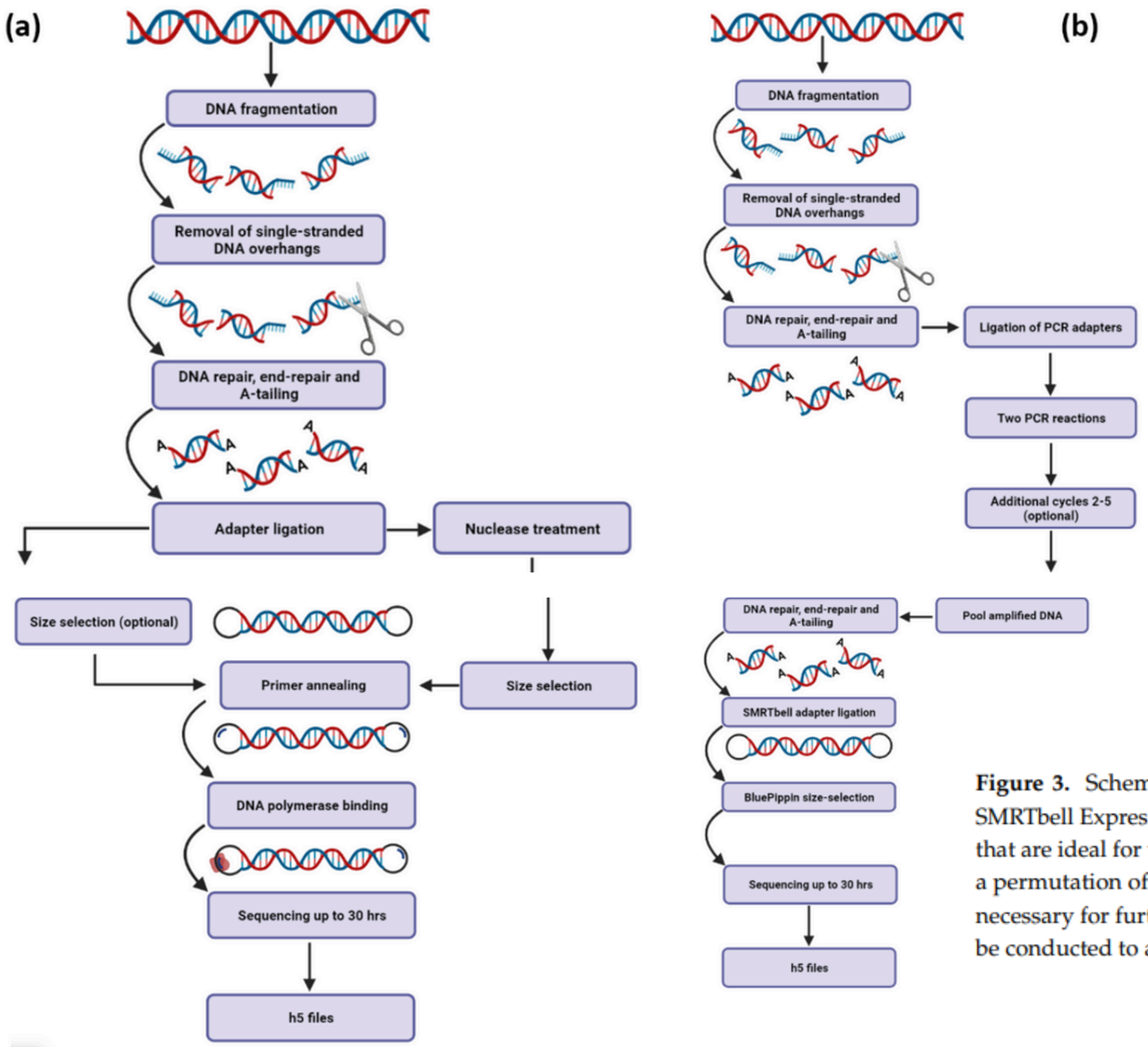
LONG READS

Sensitive to all variant types

SNVs (1 bp)	Indels (<50 bp)	Structural variants (≥50 bp)
5 Mb	3 Mb	10 Mb

Single molecule real time (SMRT)

Pacific Biosciences Single Molecule Real Time (SMRT) Sequencing



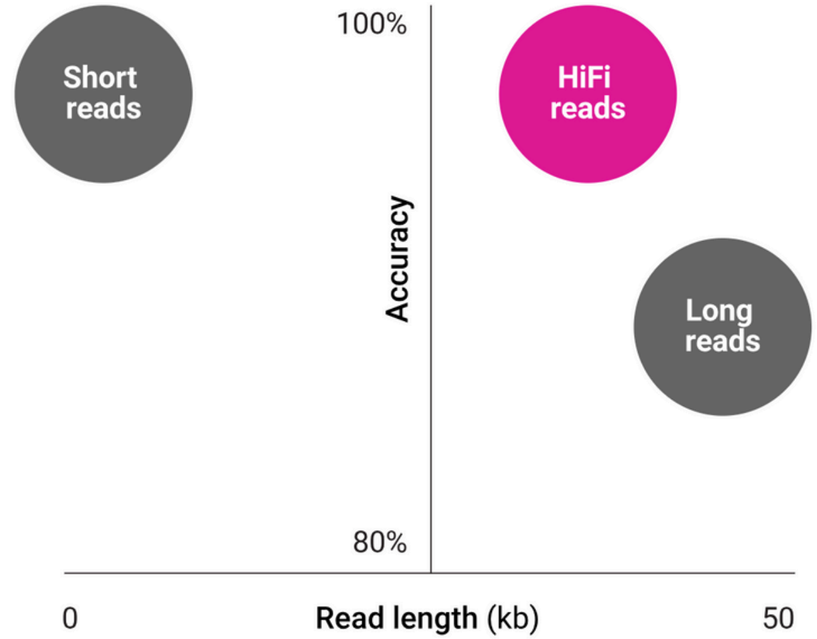
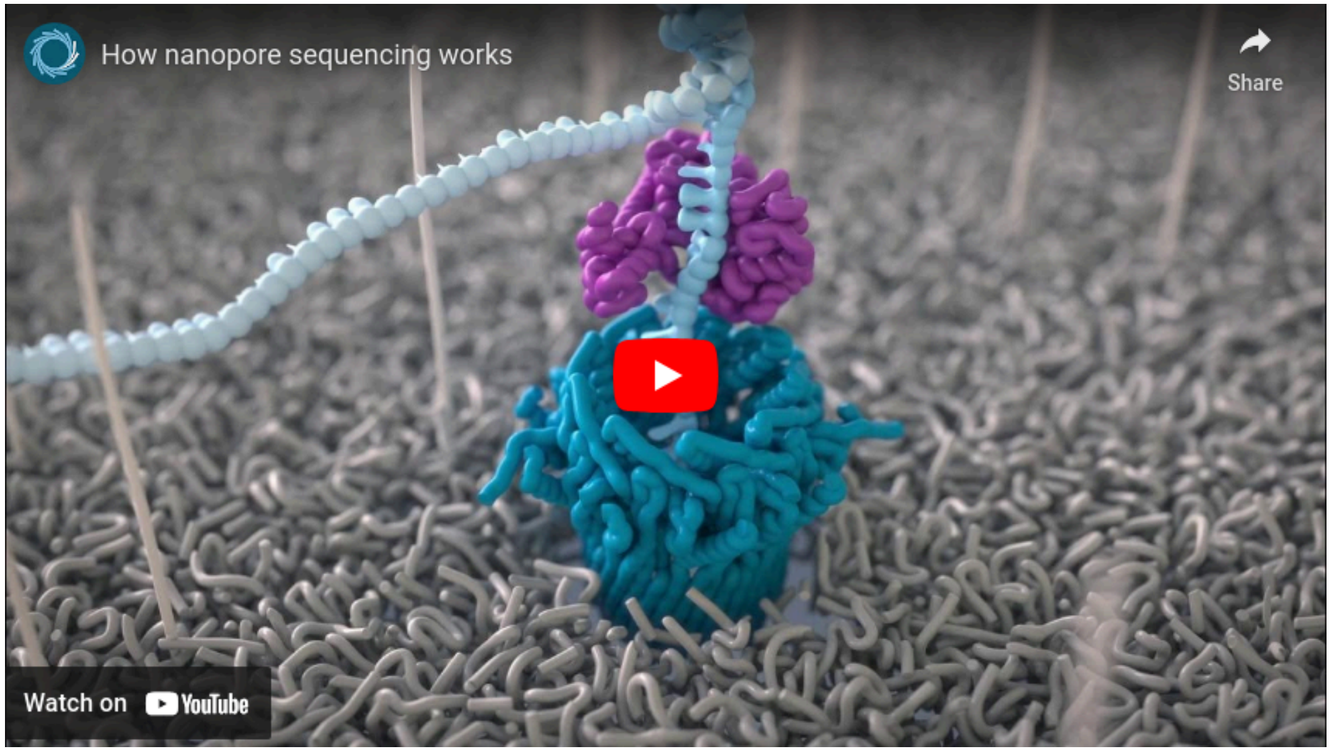


Figure 3. Schematic demonstration of PacBio's main workflows for DNA sequencing. (**a**) The SMRTbell Express approach uses gDNA as a template for the construction of HiFi SMRTbell libraries that are ideal for variant detection and de novo assembly applications. When handling larger inserts, a permutation of the approach is suggested. An additional step that includes nuclease treatment is necessary for further DNA segmentation. (**b**) In cases of ultra-low DNA inputs, PCR reactions must be conducted to achieve higher quantity of the starting material.

Oxford nanopore technology (ONT)



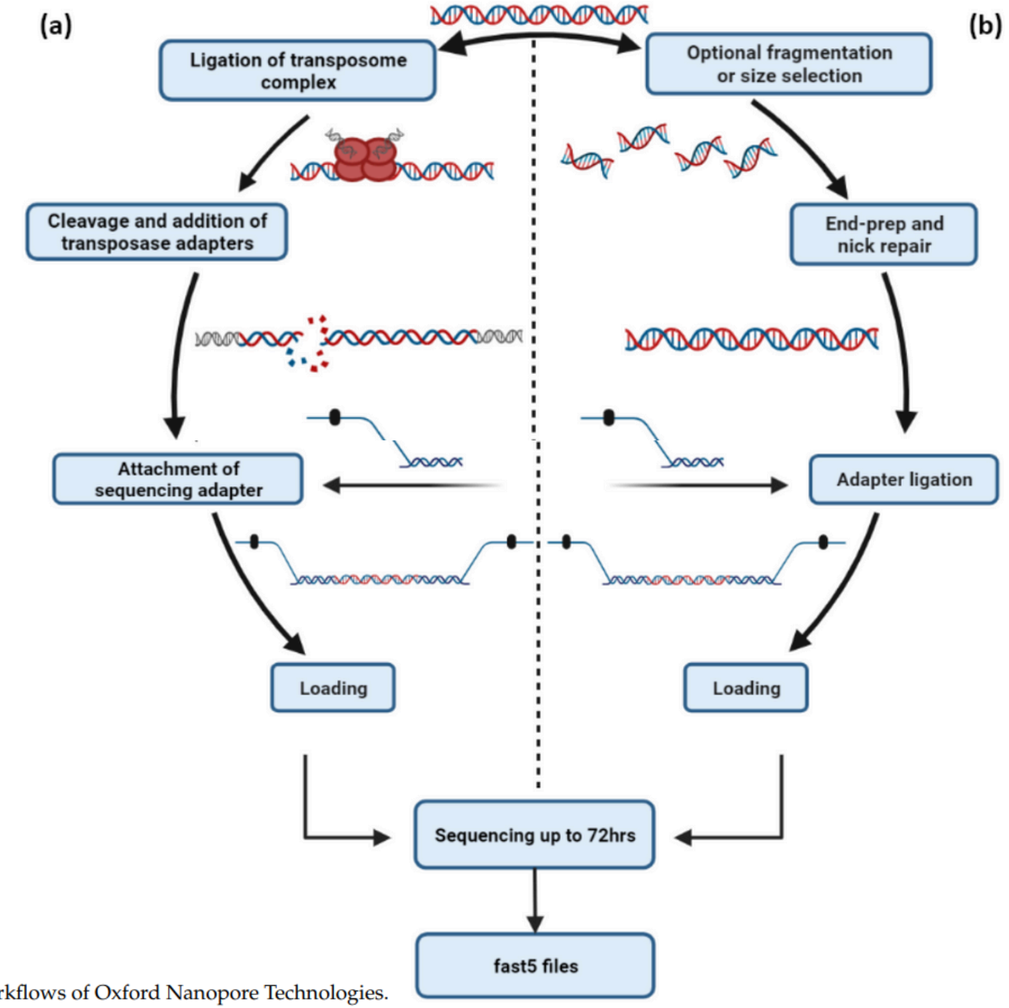


Figure 4. The most representative DNA-sequencing workflows of Oxford Nanopore Technologies. (**a**) For minimal library preparation time, ONT provides the Rapid Sequencing workflow, which exploits the innate qualities of transposase for the cleavage of genomic DNA and the subsequent adapter ligation. (**b**) For maximum throughput, ONT has developed the sequencing by ligation workflow, which includes DNA end repair and attachment of sequencing adapters for the sequencing of genomic DNA or specific amplicons.

Single molecule real time (SMRT)


PromethION 48
48
flow cells
PromethION 24
24
flow cells
PromethION™ 2
devices
2
flow cells
GridION™
5
flow cells
MinION Mk1C
1
flow cell
MinION™
Output per device
Increasing number of flow cells and / or pores

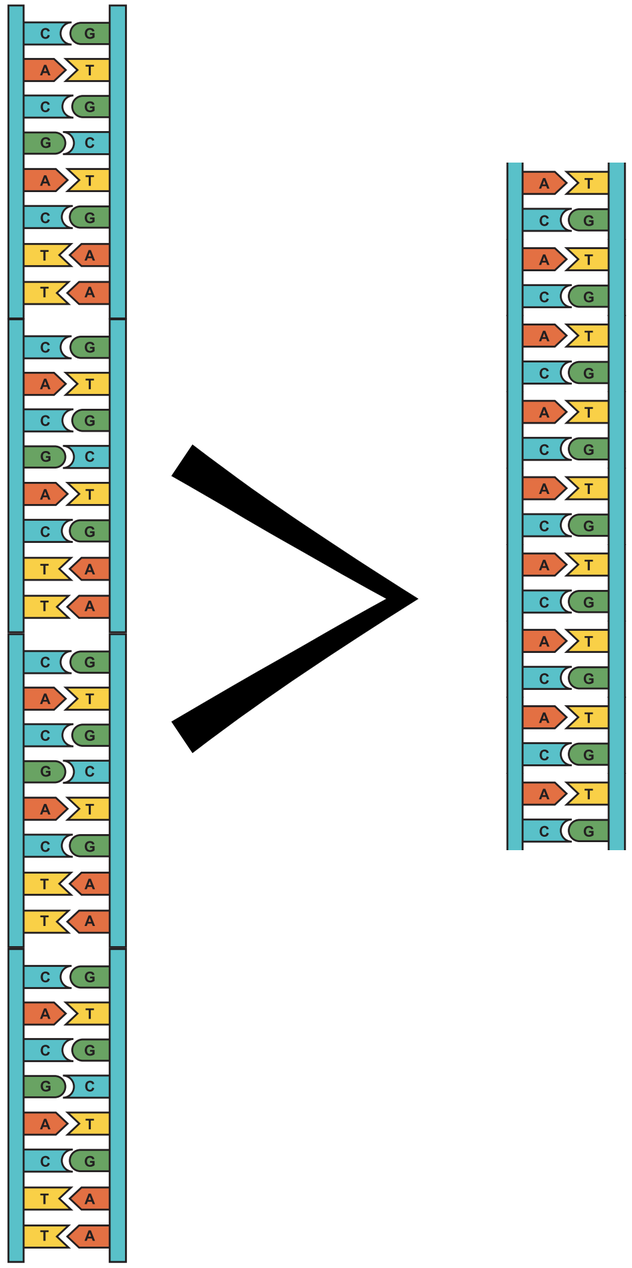
DNA Sequencing: How to Choose the Right Technology?

Montagem de Genomas

Propriedades do genoma que afetam a montagem

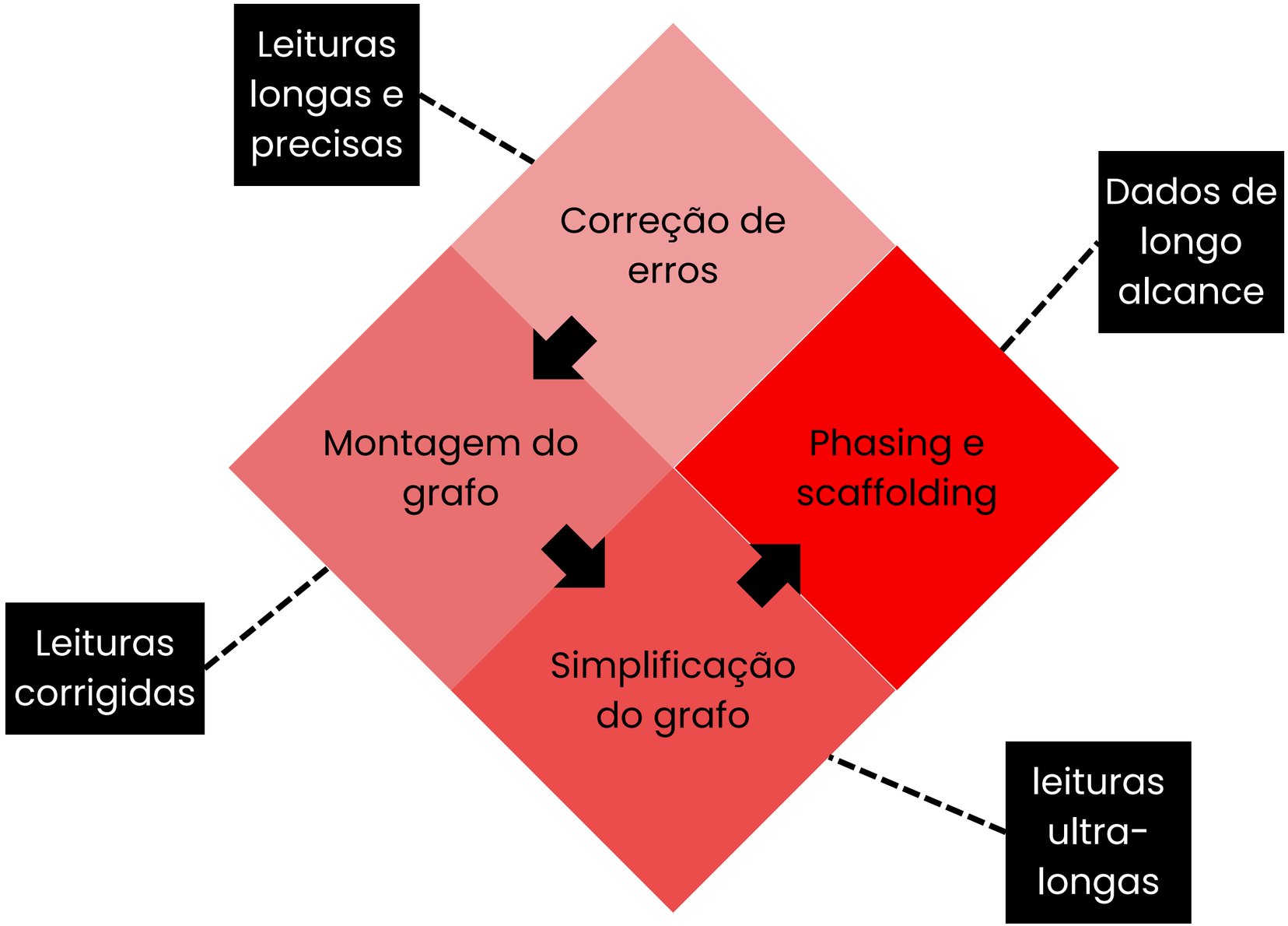
A dificuldade não é pelo tamanho e sim a estrutura repetitiva

O problema do tamanho é o requerimento computacional



ALGORITMOS DE MONTAGEM

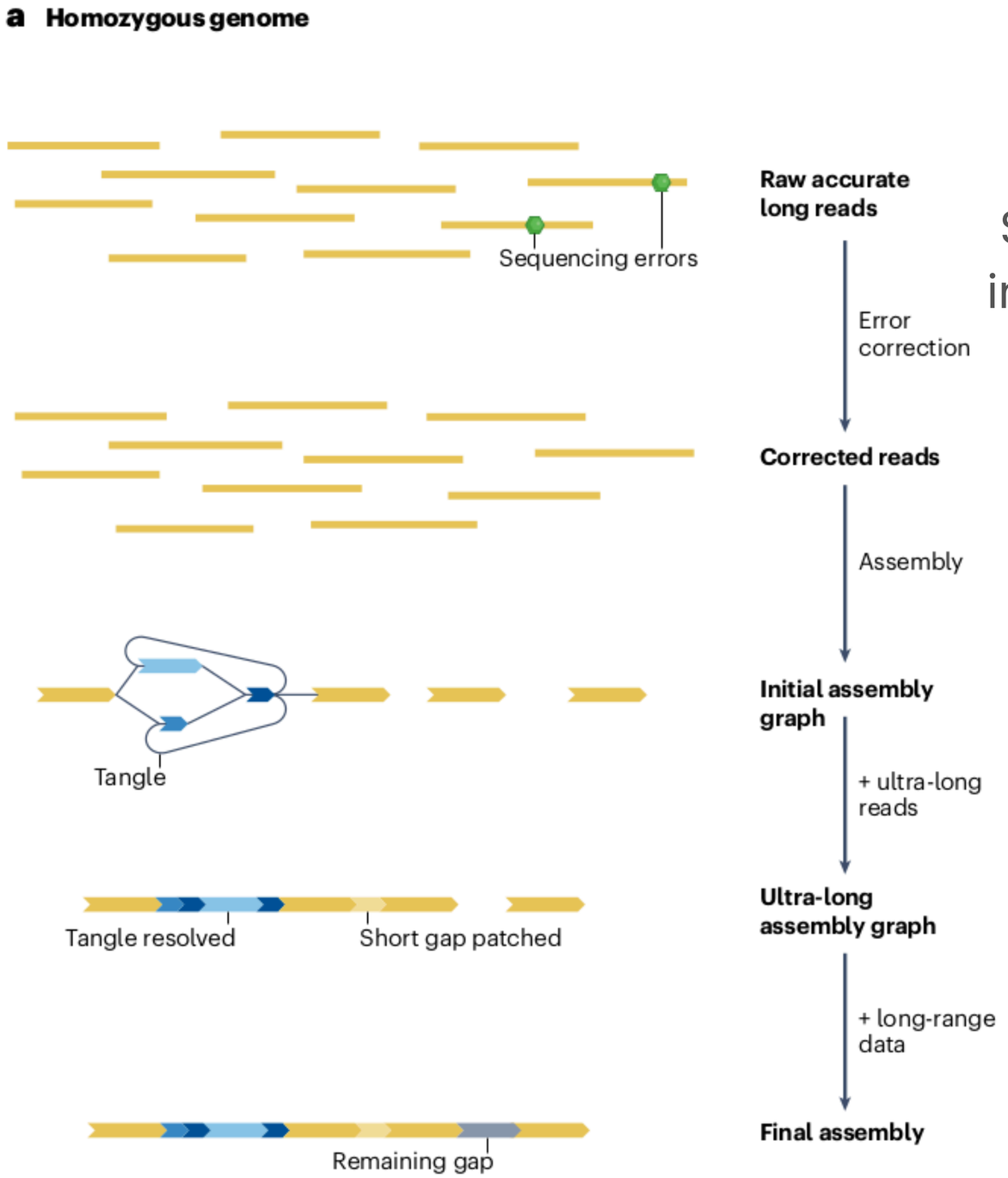
A receita atual para montagem



PARA UM GENOMA HOMOZIGOTO

Pode ser produzida com dados HiFi e ultra-longos

Um pequeno número de lacunas pode permanecer nas regiões mais desafiadoras do genoma



Softwares que integram esses dois tipos de dados

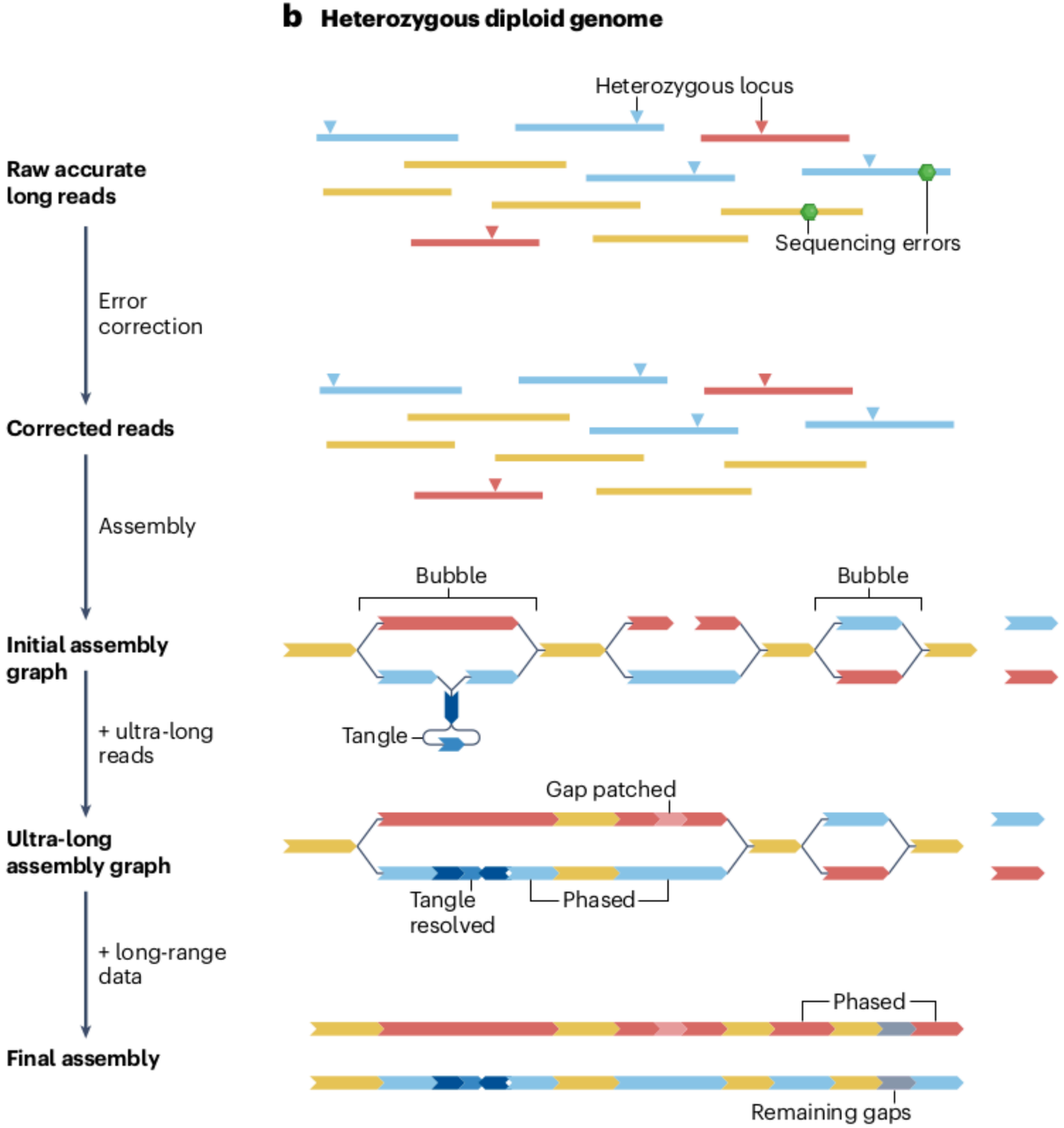
Verkko

HiFiASM

PARA UM GENOMA HETEROZIGOTO

Precisa de mais tipos de dados

Para resolver corretamente emaranhados e regiões de baixa heterozigosidade



Softwares que integram esses dois tipos de dados

Verkko

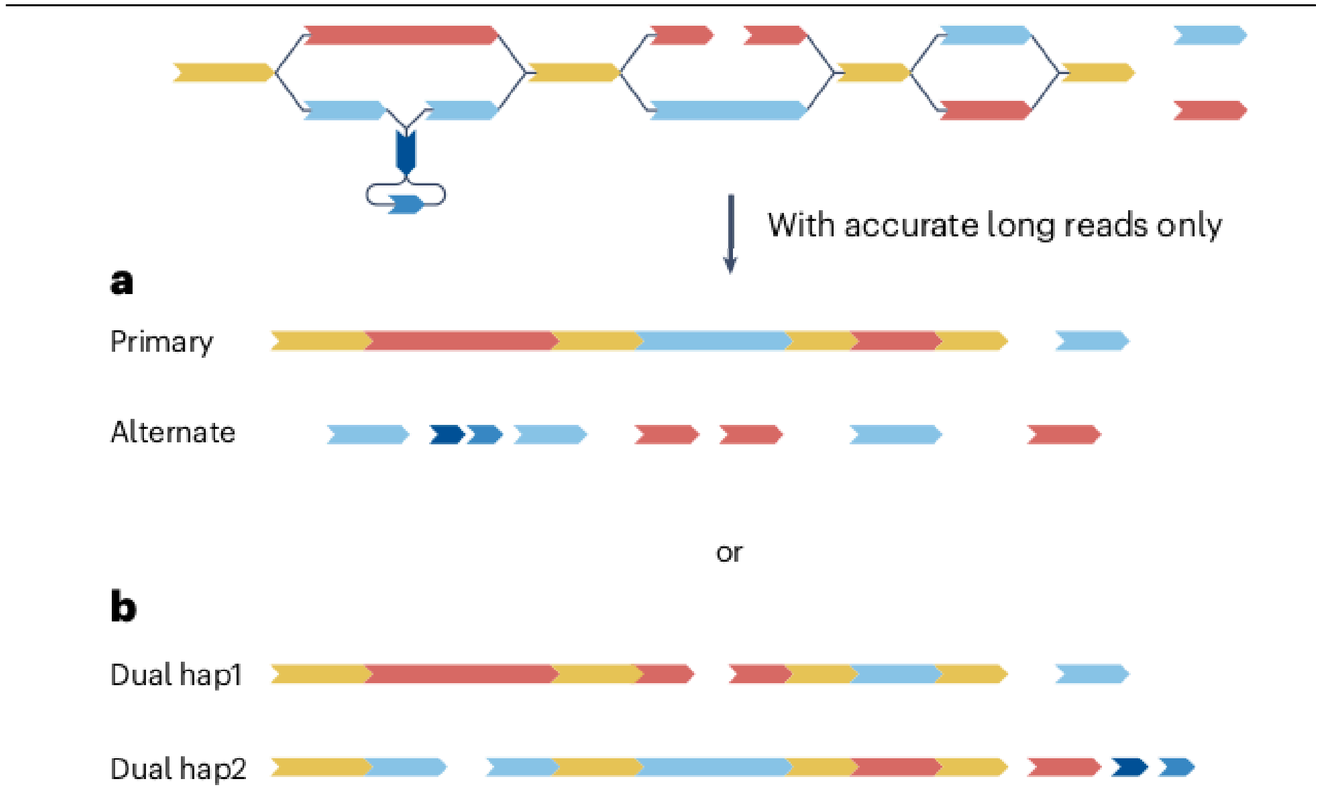
HiFiASM

PARA UM GENOMA HETEROZIGOTO

Com HiFi sozinho, podem ser produzidos dois tipos de pares de montagem:

Um par primário-alternativo

Par de montagem dupla



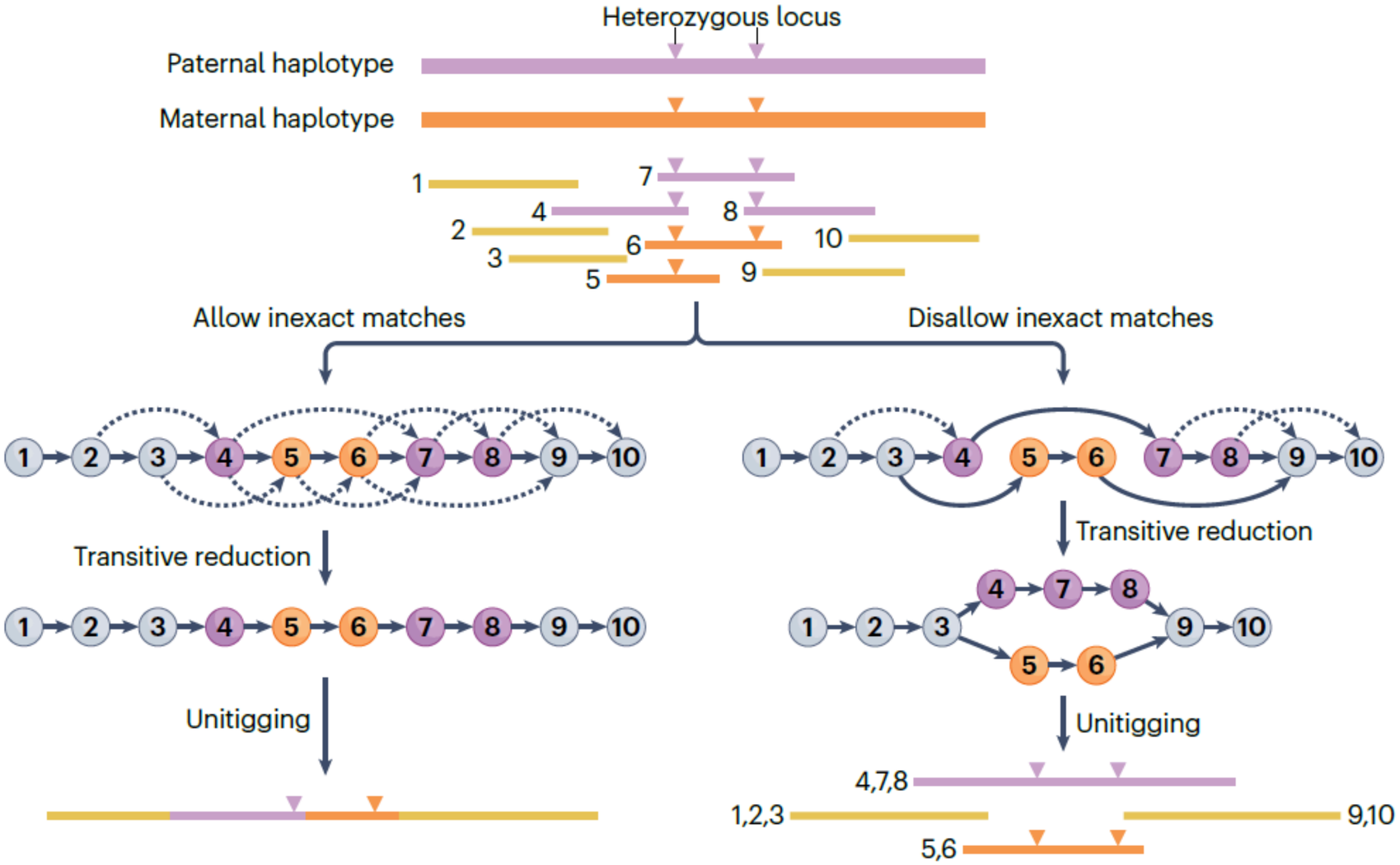
Softwares que integram esses dois tipos de dados

Verkko

HiFiASM

MONTAGEM COM TRIOS

HiFiASM



ANOTAÇÃO DO GENOMA

A primeira fase da anotação é a identificação de repetições e masking

Existem dois tipos de repetições

Sequencias de “baixa complexidade”

Elementos transponiveis

Softwares

Repeat Masker

Earl Grey

ANOTAÇÃO DO GENOMA

Anotações do genoma envolvem a caracterização de elementos de significância biológica

Principalmente na identificação de genes codificadores de proteínas

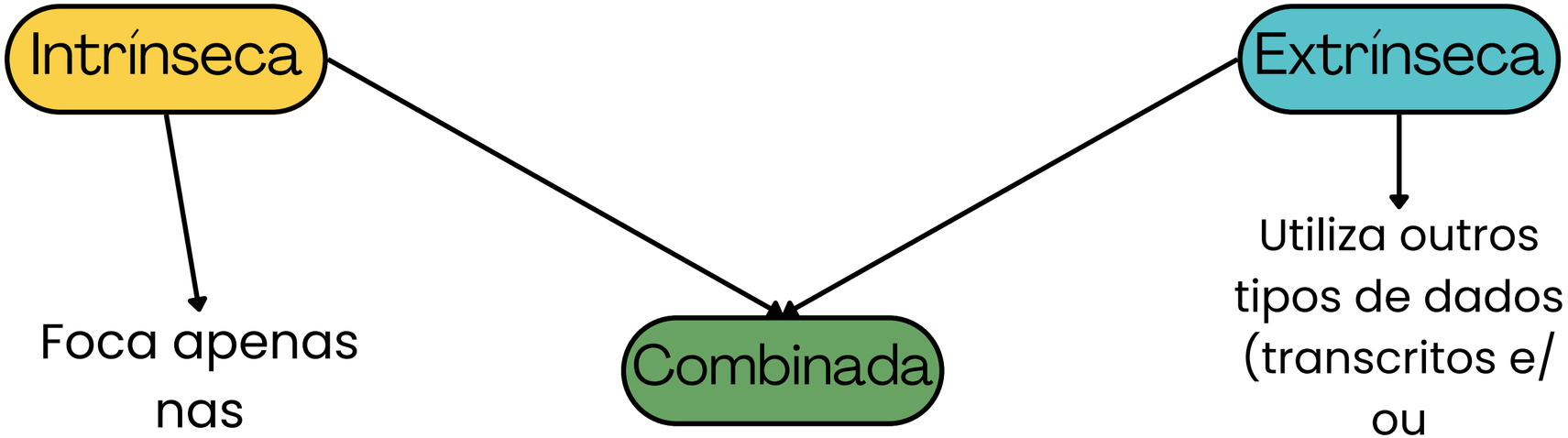
No geral, existem 3 abordagens que podem ser tomadas para predizer os genes no genoma

- Intrínseca
- Extrínseca
- Combinada

Softwares

- INFERNAL
- tRNAscan-se

ANOTAÇÃO DO GENOMA



Softwares

INFERNAL

tRNAscan-se

ANOTAÇÃO DO GENOMA

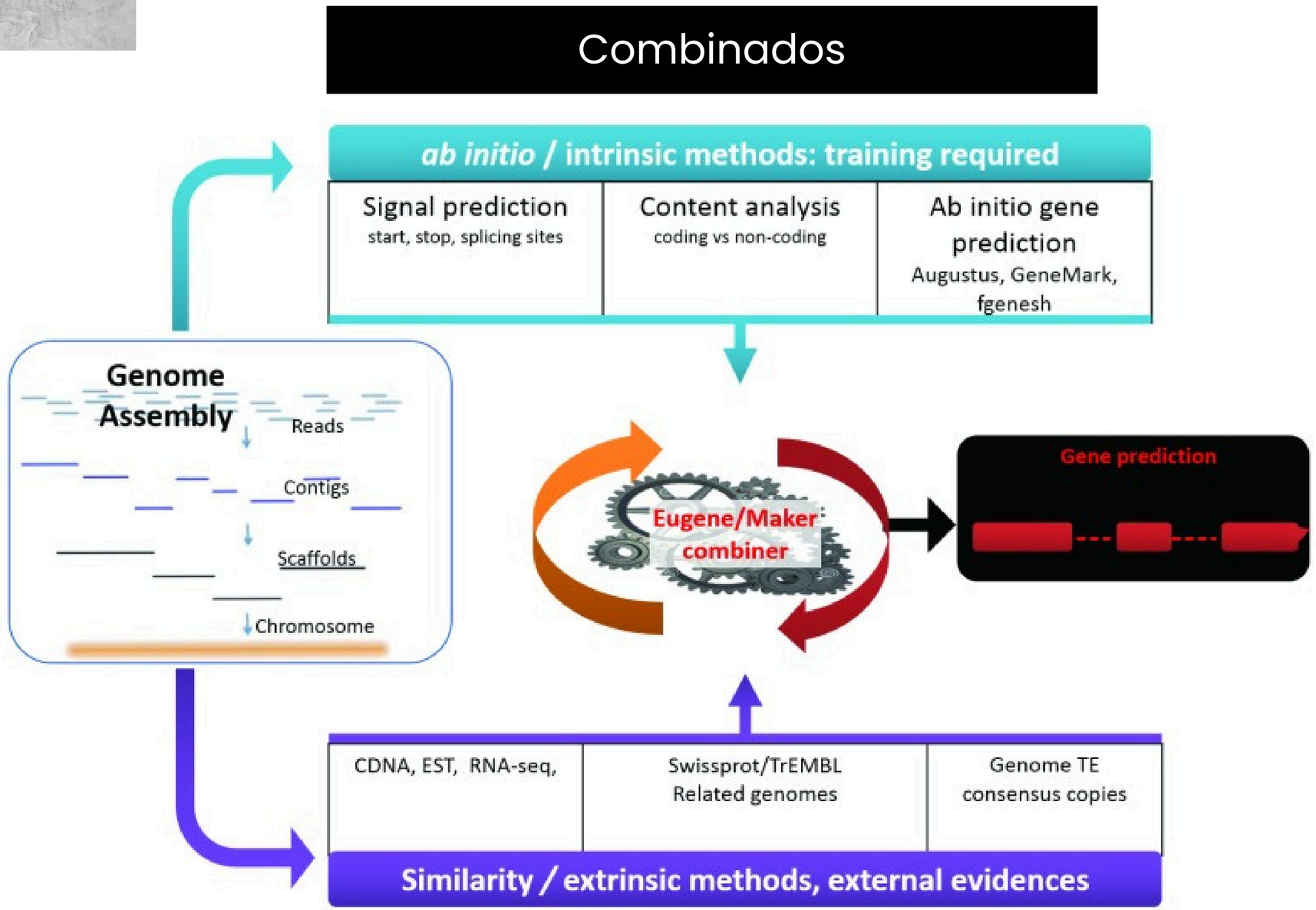
Intrínseca

É mais trabalhoso, pois modelos estatísticos precisam ser construídos e treinados.
Mas, uma grande vantagem é a capacidade de predizer genes de evolução rápida e genes únicos de espécies

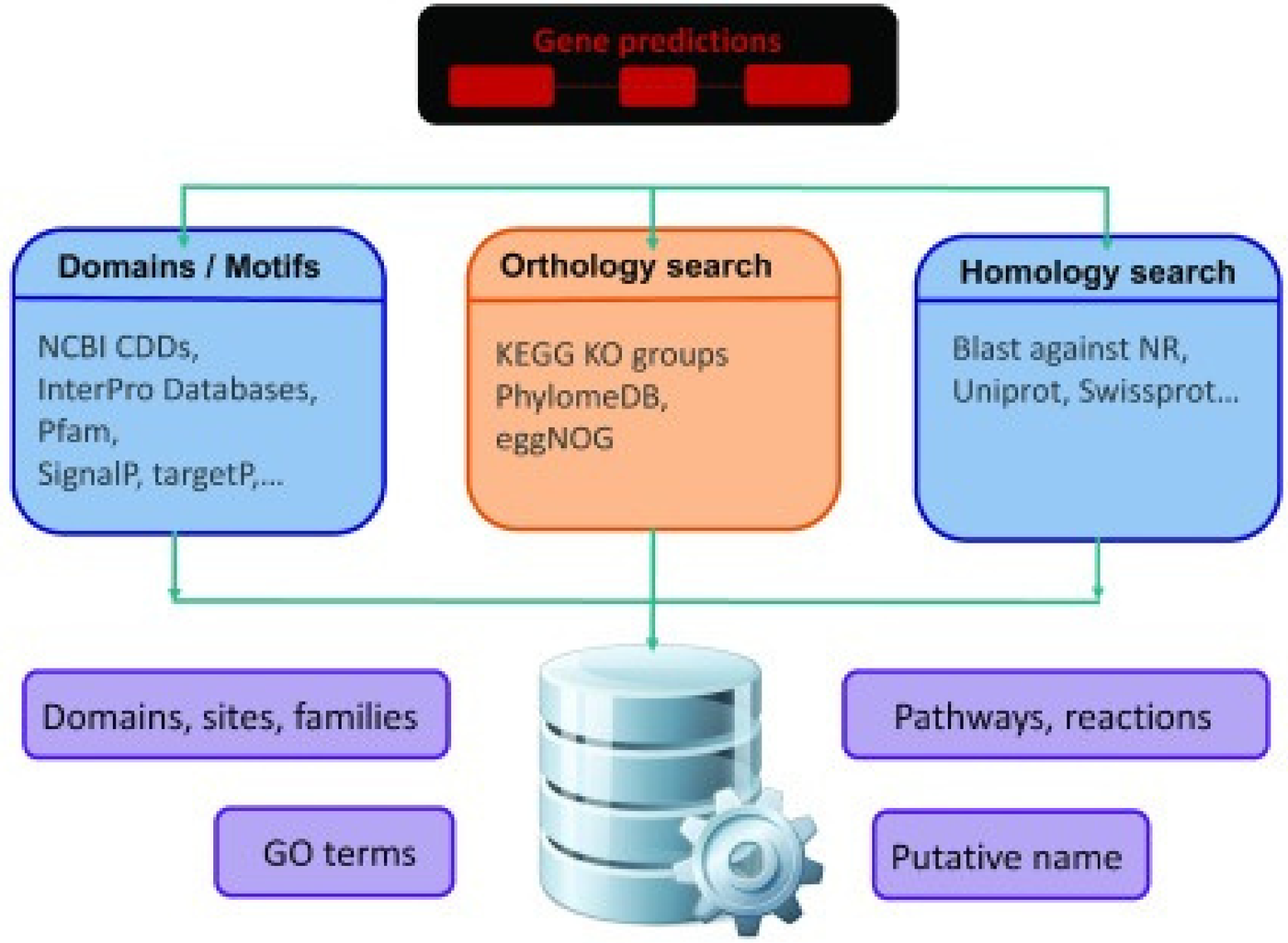
Extrínseca

É universalmente aplicável, existem vários bancos de dados com as sequências polipeptídicas (RefSeq,UniProt) que criam potencial para a predição de genes de fácil acesso

ANOTAÇÃO DO GENOMA



ANOTAÇÃO FUNCIONAL

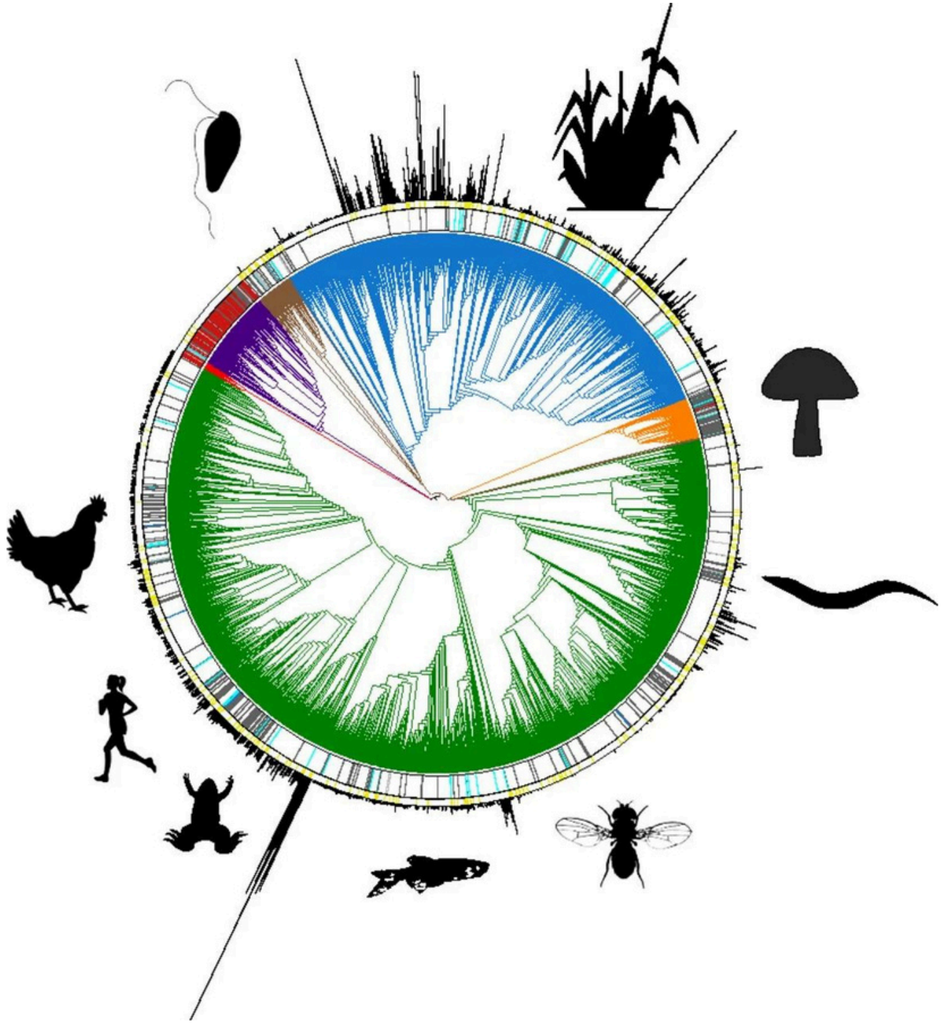


INICIATIVAS E PROGRAMAS DE GENOMICA

PERSPECTIVE | BIOLOGICAL SCIENCES |

Earth BioGenome Project: Sequencing life for the future of life

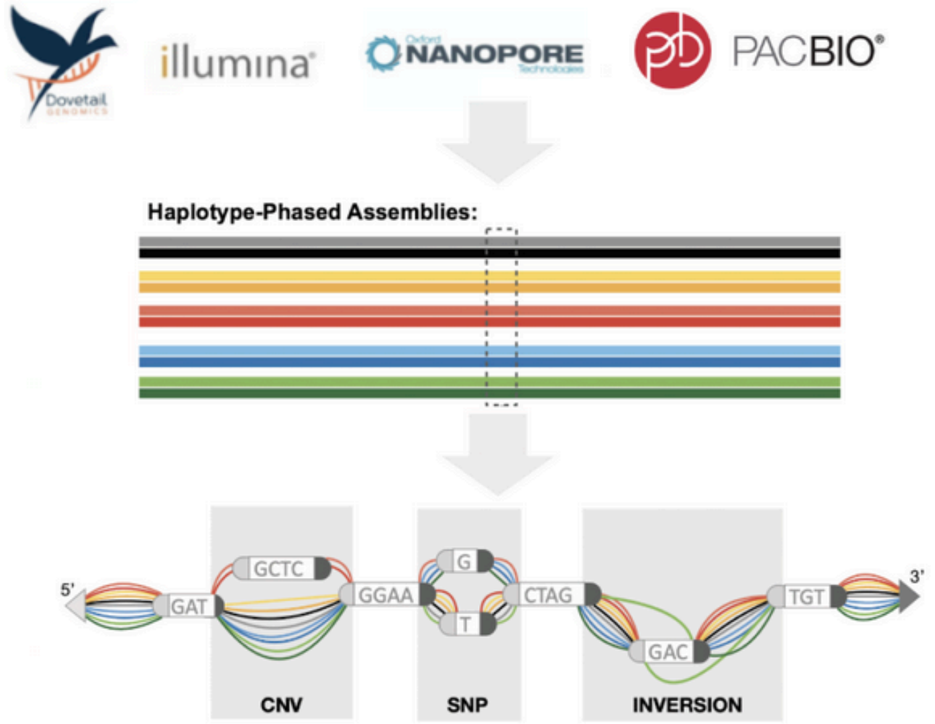
Aumentar o nosso entendimento da biodiversidade da Terra e responsavelmente direcionar os recursos são um dos maiores desafios da ciência e da sociedade no novo milênio



INICIATIVAS E PROGRAMAS DE GENOMICA

Human Pangenome Reference Consortium

É um projeto fundado para sequenciar e montar os genomas de indivíduos de diversas populações para melhor representar a diversidade genômica na população humana



REFERÊNCIAS BIBLIOGRÁFICAS

1. Athanasopoulou, K., Boti, M. A., Adamopoulos, P. G., Skourou, P. C., & Scorilas, A. (2021). Third-Generation Sequencing: The Spearhead towards the Radical Transformation of Modern Genomics. Life, 12(1), 30. https://doi.org/10.3390/life12010030
2. Cooper, G. M. (2000). The complexity of eukaryotic genomes. The Cell - NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/NBK9846/
3. Discovery of DNA double helix: Watson and Crick | Learn Science at Scitable. (n.d.). https://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397/
4. Gasc, C., Ribière, C., Parisot, N., Beugnot, R., Defois, C., Petit-Biderre, C., Boucher, D., Peyretaillade, E., & Peyret, P. (2015). Capturing prokaryotic dark matter genomes. Research in Microbiology, 166(10), 814–830. https://doi.org/10.1016/j.resmic.2015.06.001
5. Kumar, K. R., Cowley, M. J., & Davis, R. L. (2024). Next-Generation Sequencing and Emerging Technologies*. Seminars in Thrombosis and Hemostasis. https://doi.org/10.1055/s-0044-1786397
6. Mardis, E. R. (2013). Next-Generation Sequencing Platforms. Annual Review of Analytical Chemistry, 6(1), 287–303. https://doi.org/10.1146/annurev-anchem-062012-092628
7. Maxam, A. M., & Gilbert, W. (1977). A new method for sequencing DNA. Proceedings of the National Academy of Sciences, 74(2), 560–564. https://doi.org/10.1073/pnas.74.2.560
8. Mobley, I. (2024, July 26). DNA sequencing: How to choose the right technology - Front line genomics. Front Line Genomics. https://frontlinegenomics.com/dna-sequencing-how-to-choose-the-right-technology/
9. Nhgri. (2019, March 9). A brief guide to genomics. Genome.gov. https://www.genome.gov/about-genomics/fact-sheets/A-Brief-Guide-to-Genomics
10. PacBio. (2024, August 21). Whole genome sequencing - PacBio. https://www.pacb.com/products-and-services/applications/whole-genome-sequencing/
11. Rusk, N. (2009). Cheap third-generation sequencing. Nature Methods, 6(4), 244. https://doi.org/10.1038/nmeth0409-244a
12. Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences, 74(12), 5463–5467. https://doi.org/10.1073/pnas.74.12.5463
13. Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., & Waterston, R. H. (2017). DNA sequencing at 40: past, present and future. Nature, 550(7676), 345–353. https://doi.org/10.1038/nature24286
14. Soppa, J. (2021). Non-equivalent genomes in polyploid prokaryotes. Nature Microbiology, 7(2), 186–188. https://doi.org/10.1038/s41564-021-01034-3
15. The Editors of Encyclopaedia Britannica. (2024, September 17). Genetic engineering | Definition, Process, Uses, Examples, Techniques, & Facts. Encyclopedia Britannica. https://www.britannica.com/science/genetic-engineering/Process-and-techniques
16. Timeline: History of Genomics. (n.d.). https://www.yourgenome.org/theme/timeline-history-of-genomics/
17. Wang, Y., Zhao, Y., Bollas, A., Wang, Y., & Au, K. F. (2021). Nanopore sequencing technology, bioinformatics and applications. Nature Biotechnology, 39(11), 1348–1365. https://doi.org/10.1038/s41587-021-01108-x
18. What's in a genome? (n.d.). Earlham Institute. https://www.earlham.ac.uk/articles/whats-genome