Estrutura e Função do DNA

IBM-1027 - Genética Molecular

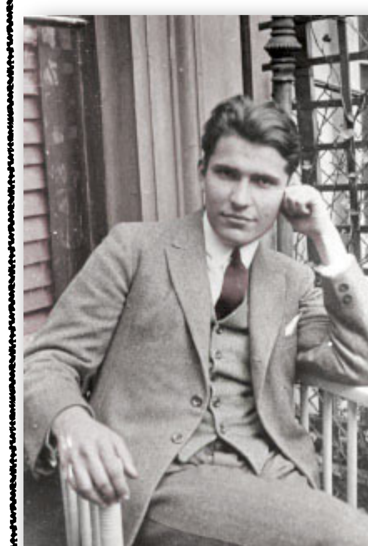
Wilson Araújo Silva Jr
FMRP/USP

Friedrich Miescher
Isolou o “DNA” pela primeira vez (nucleína).

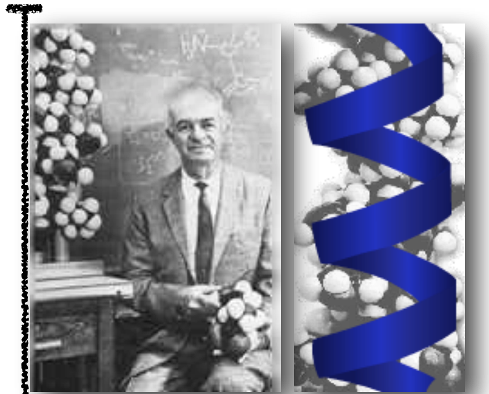
Oswald Theodore Avery
O DNA carregava a informação genética.

1869 1928 1944 1950

1928 - Fredrick Griffith
Princípio Transformante



Erwin Chargaff
A relação de bases A=T e C=G é constante em DNA das mais variadas espécies.



L. Pauling

Conceito de alfa Hélice (Proteínas) e da natureza das ligações químicas. A ponte de hidrogênio tem 3% da força da ligação covalente

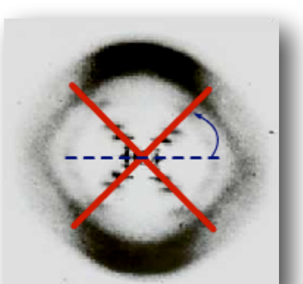
Hershey-Chase

Confirmação de que O DNA carregava a informação genética.

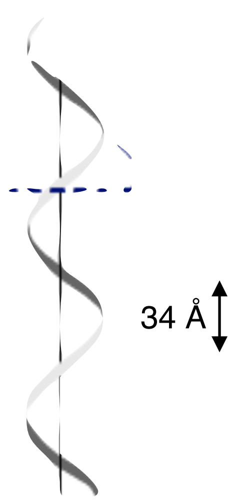
1951 1952 1953

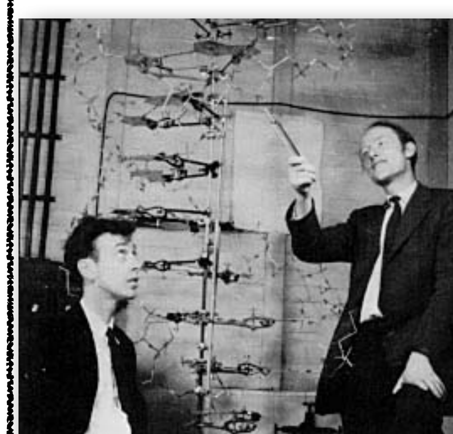
Rosalin Franklin e Maurice Wilkins

Hélice com periodicidade regular de 3,4 Å e 34 Å



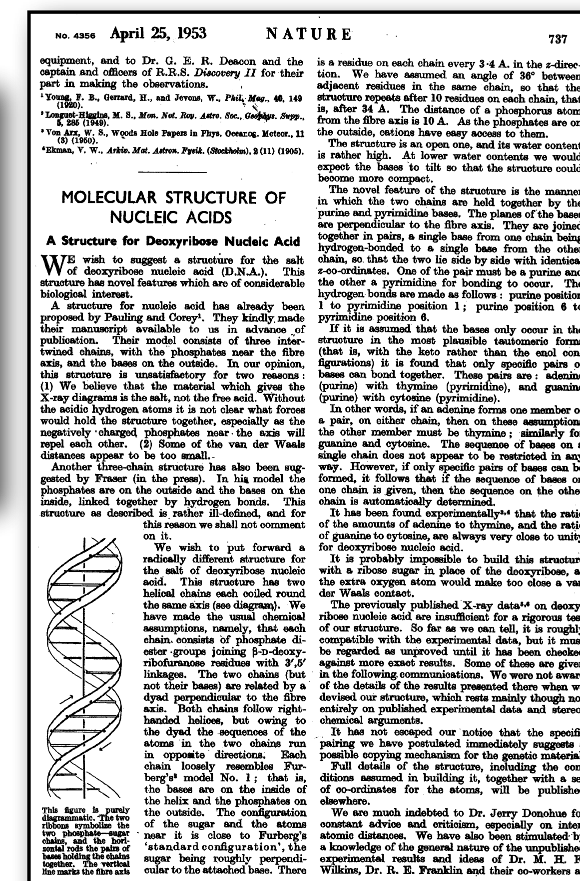
Difração de Raios-X





Watson e Crick

Estrutura molecular do DNA



No. 4356 April 25, 1953 NATURE 737

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, 40, 149 (1920).
[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, 5, 285 (1949).
[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., 11 (3) (1950).
[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik. (Stockholm)*, 2 (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three intertwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

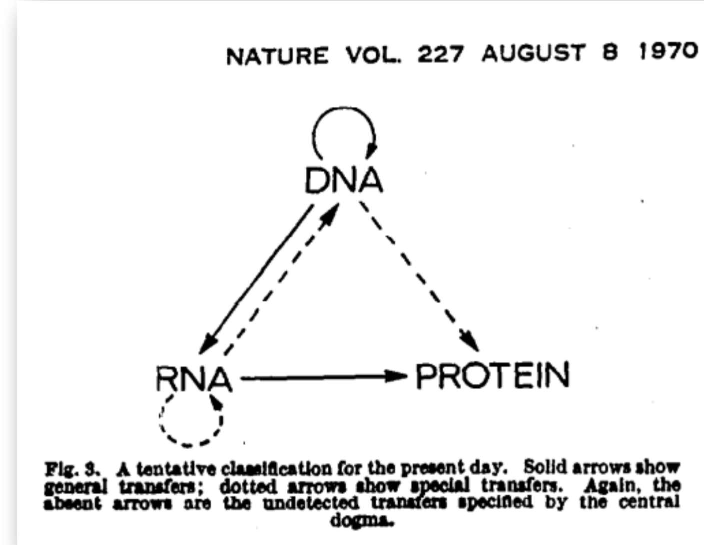
It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

Matthew Meselson e Franklin W. Stahl

Modelo de replicação do DNA semiconservativa.



Dogma Central da Biologia Molecular

Watson e Crick

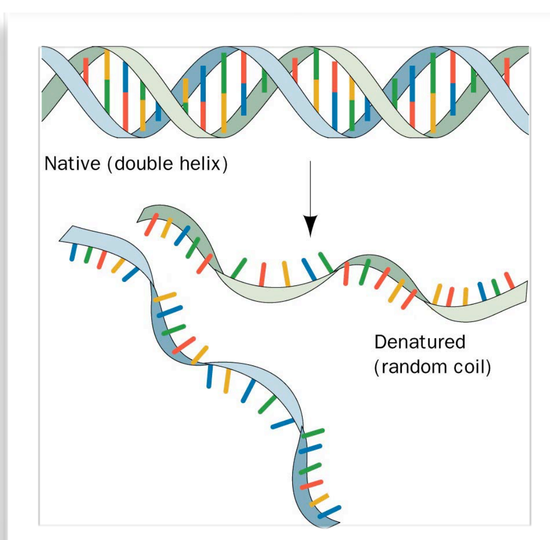
Postulado em 1958

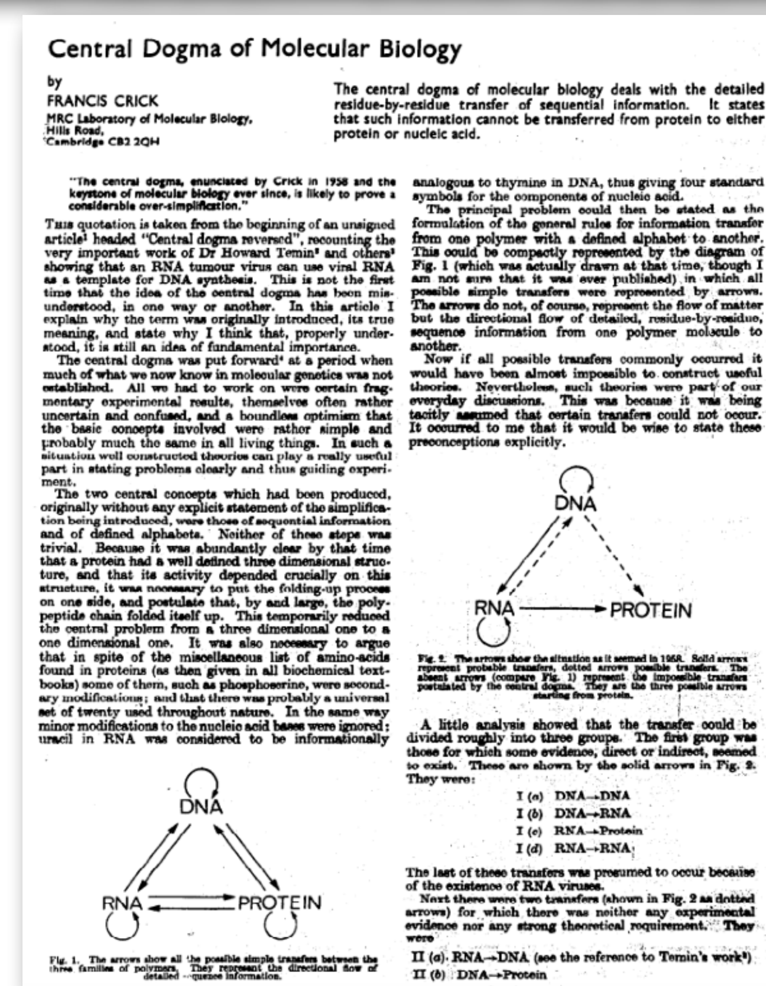
Publicado em 1970

1958 1961 1970

Marmur e Doty

Renaturação de DNA





Central Dogma of Molecular Biology

by
FRANCIS CRICK
MRC Laboratory of Molecular Biology,
Hills Road,
Cambridge CB2 2QH

The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein or nucleic acid.

"The central dogma, enunciated by Crick in 1958 and the keystone of molecular biology ever since, is likely to prove a considerable over-simplification."

THIS quotation is taken from the beginning of an unsigned article[1] headed "Central dogma reversed", recounting the very important work of Dr Howard Temin[2] and others[3] showing that an RNA tumour virus can use viral RNA as a template for DNA synthesis. This is not the first time that the idea of the central dogma has been misunderstood, in one way or another. In this article I explain why the term was originally introduced, its true meaning, and state why I think that, properly understood, it is still an idea of fundamental importance.

The central dogma was put forward[4] at a period when much of what we now know in molecular genetics was not established. All we had to work on were certain fragmentary experimental results, themselves often rather uncertain and confused, and a boundless optimism that the basic concepts involved were rather simple and probably much the same in all living things. In such a situation well constructed theories can play a really useful part in stating problems clearly and thus guiding experiment.

The two central concepts which had been produced, originally without any explicit statement of the simplification being introduced, were those of sequential information and of defined alphabets. Neither of these steps was trivial. Because it was abundantly clear by that time that a protein had a well defined three dimensional structure, and that its activity depended crucially on this structure, it was necessary to put the folding-up process on one side, and postulate that, by and large, the polypeptide chain folded itself up. This temporarily reduced the central problem from a three dimensional one to a one dimensional one. It was also necessary to argue that in spite of the miscellaneous list of amino-acids found in proteins (as then given in all biochemical textbooks) some of them, such as phosphoserine, were secondary modifications; and that there was probably a universal set of twenty used throughout nature. In the same way minor modifications to the nucleic acid bases were ignored; uracil in RNA was considered to be informationally analogous to thymine in DNA, thus giving four standard symbols for the components of nucleic acid.

The principal problem could then be stated as the formulation of the general rules for information transfer from one polymer with a defined alphabet to another. This could be compactly represented by the diagram of Fig. 1 (which was actually drawn at that time, though I am not sure that it was ever published) in which all possible simple transfers were represented by arrows. The arrows do not, of course, represent the flow of matter but the directional flow of detailed, residue-by-residue, sequence information from one polymer molecule to another.

Now if all possible transfers commonly occurred it would have been almost impossible to construct useful theories. Nevertheless, such theories were part of our everyday discussions. This was because it was being tacitly assumed that certain transfers could not occur. It occurred to me that it would be wise to state these preconceptions explicitly.

DNA
RNA
PROTEIN

Fig. 1. The arrows show all the possible simple transfers between the three families of polymers. They represent the directional flow of detailed sequence information.

DNA
RNA
PROTEIN

Fig. 2. The arrows show the situation as it seemed in 1958. Solid arrows represent probable transfers, dotted arrows possible transfers. The absent arrows (compare Fig. 1) represent the impossible transfers postulated by the central dogma. They are the three possible arrows starting from protein.

A little analysis showed that the transfer could be divided roughly into three groups. The first group was those for which some evidence, direct or indirect, seemed to exist. These are shown by the solid arrows in Fig. 2. They were:

I (a) DNA→DNA
I (b) DNA→RNA
I (c) RNA→Protein
I (d) RNA→RNA

The last of these transfers was presumed to occur because of the existence of RNA viruses.

Next there were two transfers (shown in Fig. 2 as dotted arrows) for which there was neither any experimental evidence nor any strong theoretical requirement. They were

II (a) RNA→DNA (see the reference to Temin's work[2])
II (b) DNA→Protein

1928 - Fredrick Griffith

Princípio Transformante

"Fred Griffith discovers the phenomenon of transformation, in which some unknown "principle" transforms a harmless strain of bacteria into a virulent one."

1944 - Oswald Theodore Avery

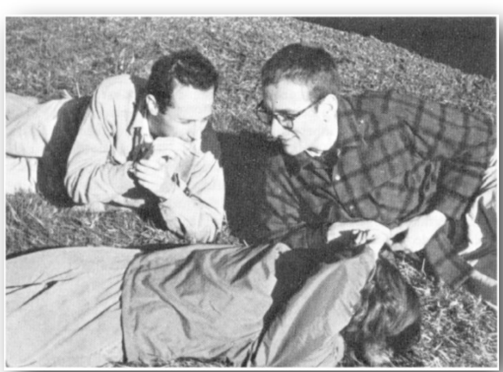
O DNA carregava a informação genética.

"Oswald Avery, Colin Macleod, and Maclyn McCarty prove that DNA, not protein, embodies the heredity material in most living organisms."

1952 - Hershey-Chase Experiments

Confirmação de que O DNA carregava a informação genética.

"Alfred Hershey and Martha Chase that helped to confirm that DNA is the genetic material."

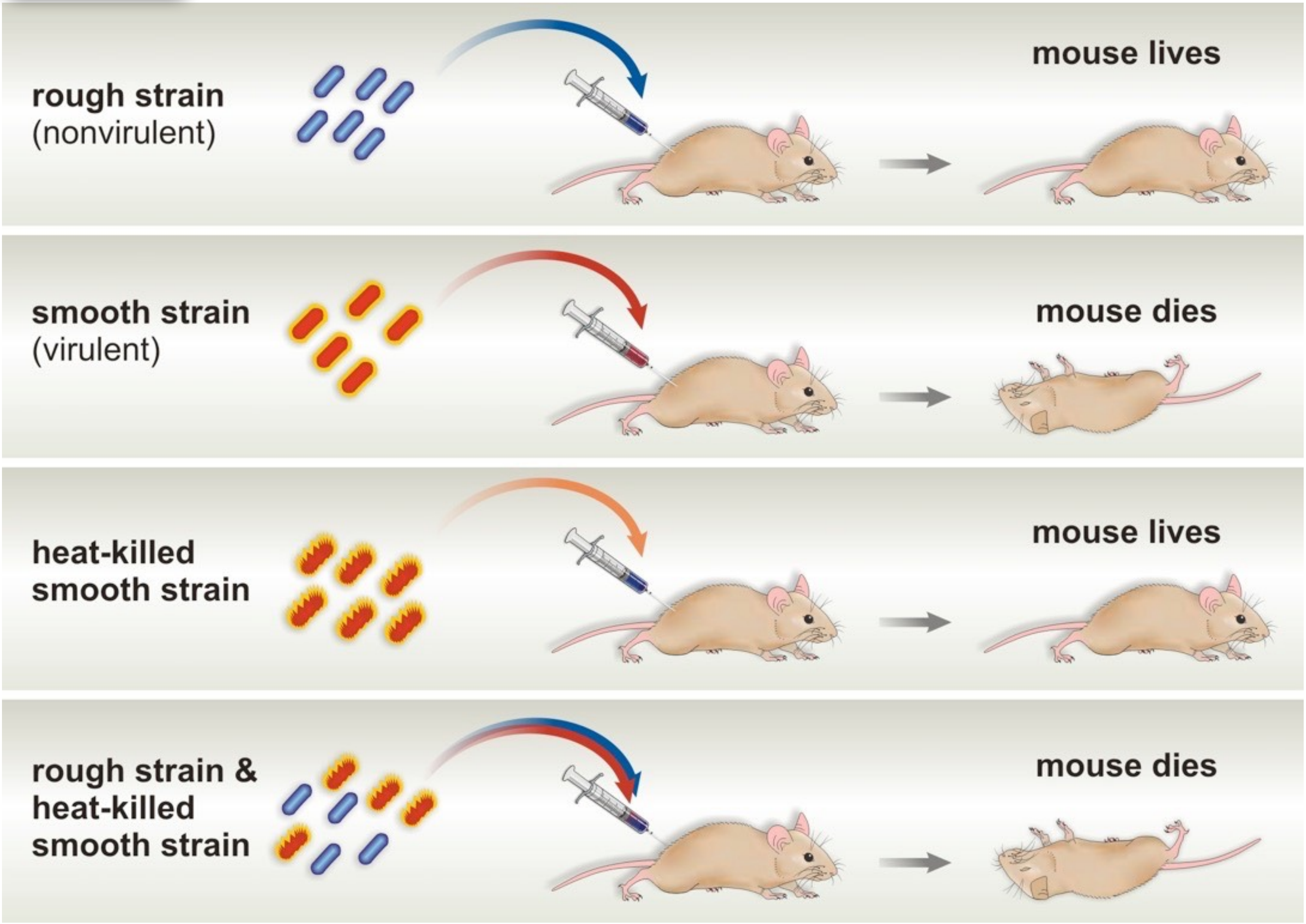


1958 - Matthew Meselson e Franklin W. Stahl

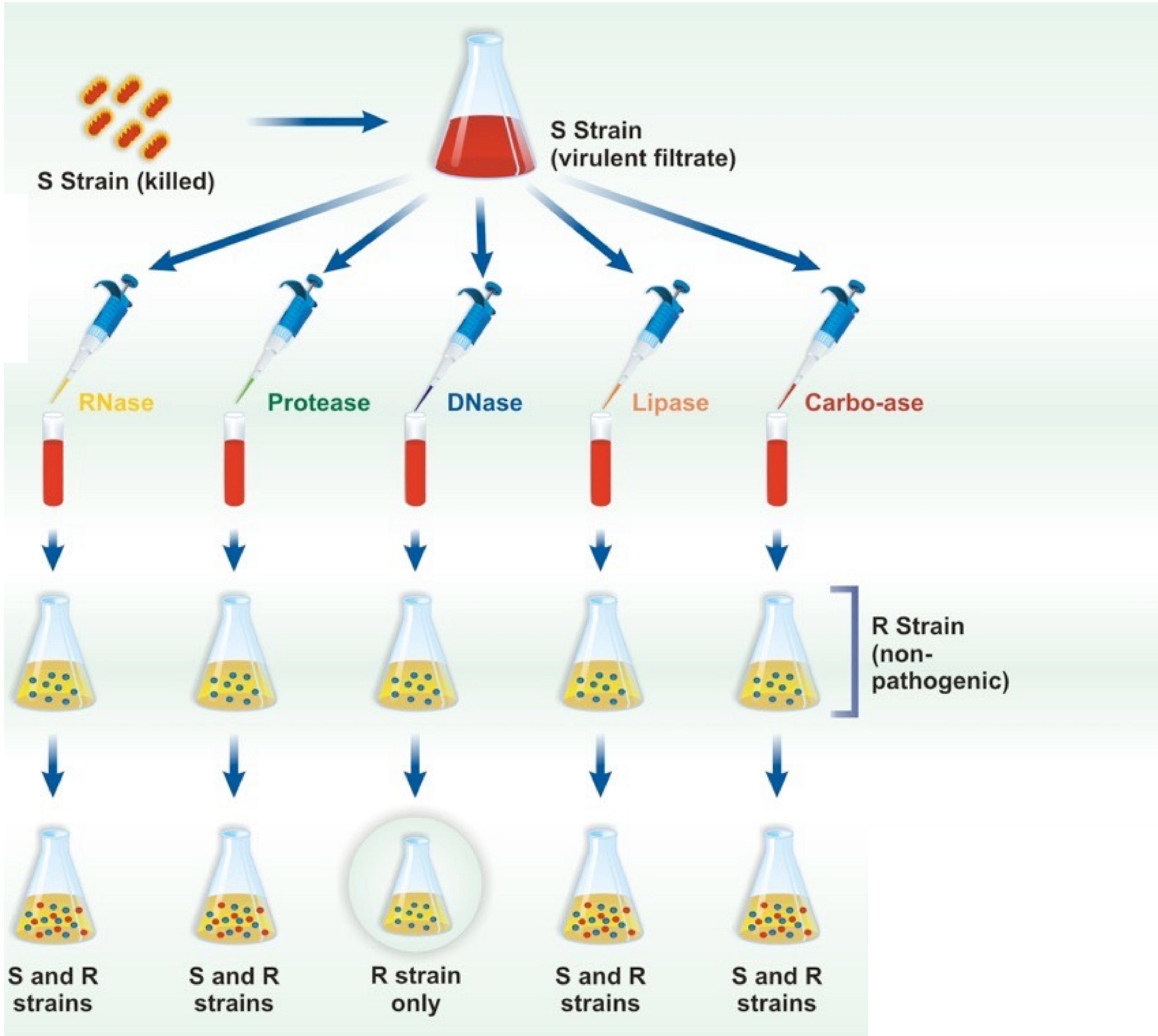
Modelo de replicação do DNA semiconservativa.

"Alfred Hershey and Martha Chase that helped to confirm that DNA is the genetic material."

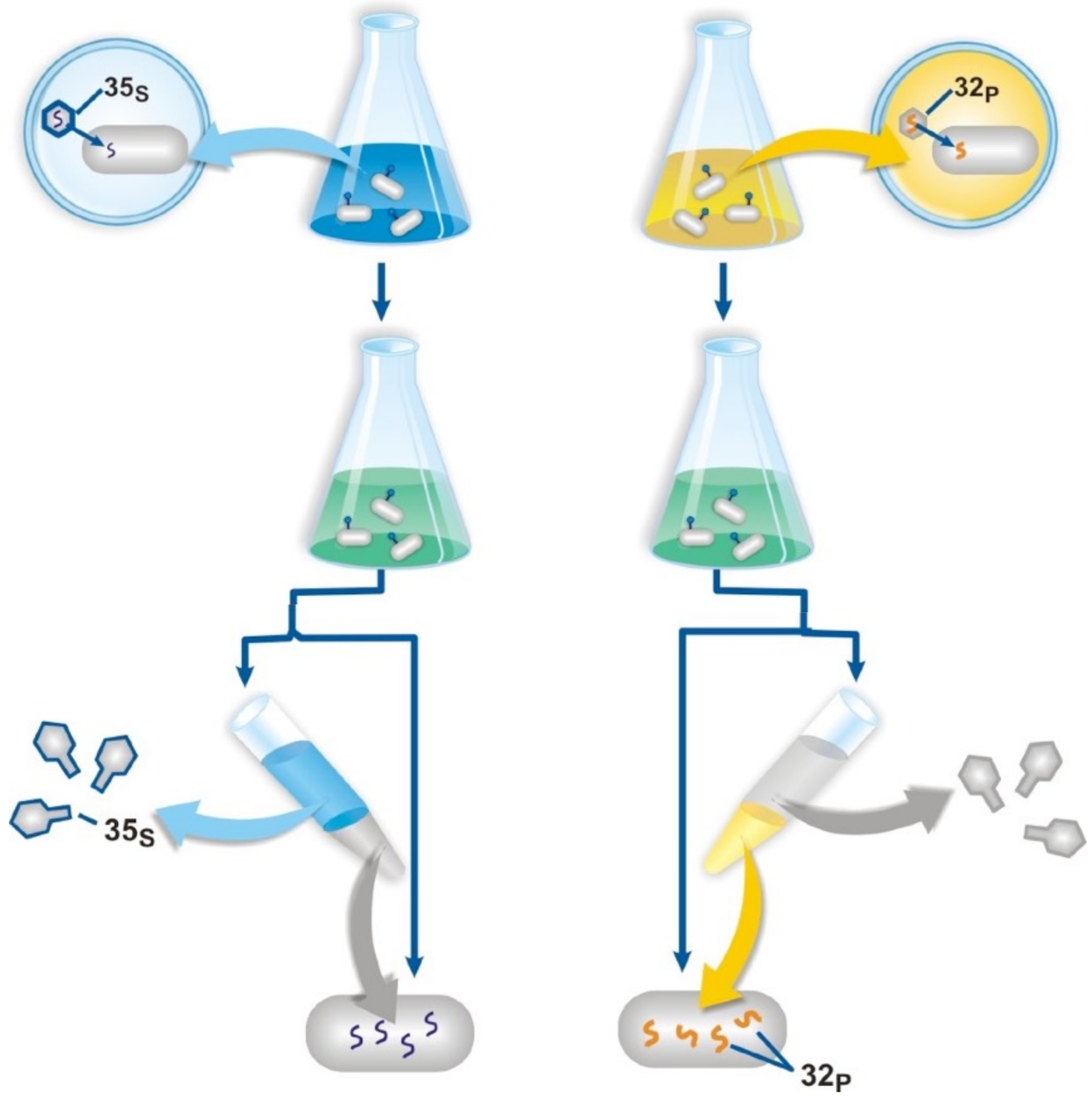
1928 - Fredrick Griffith
Princípio Transformante



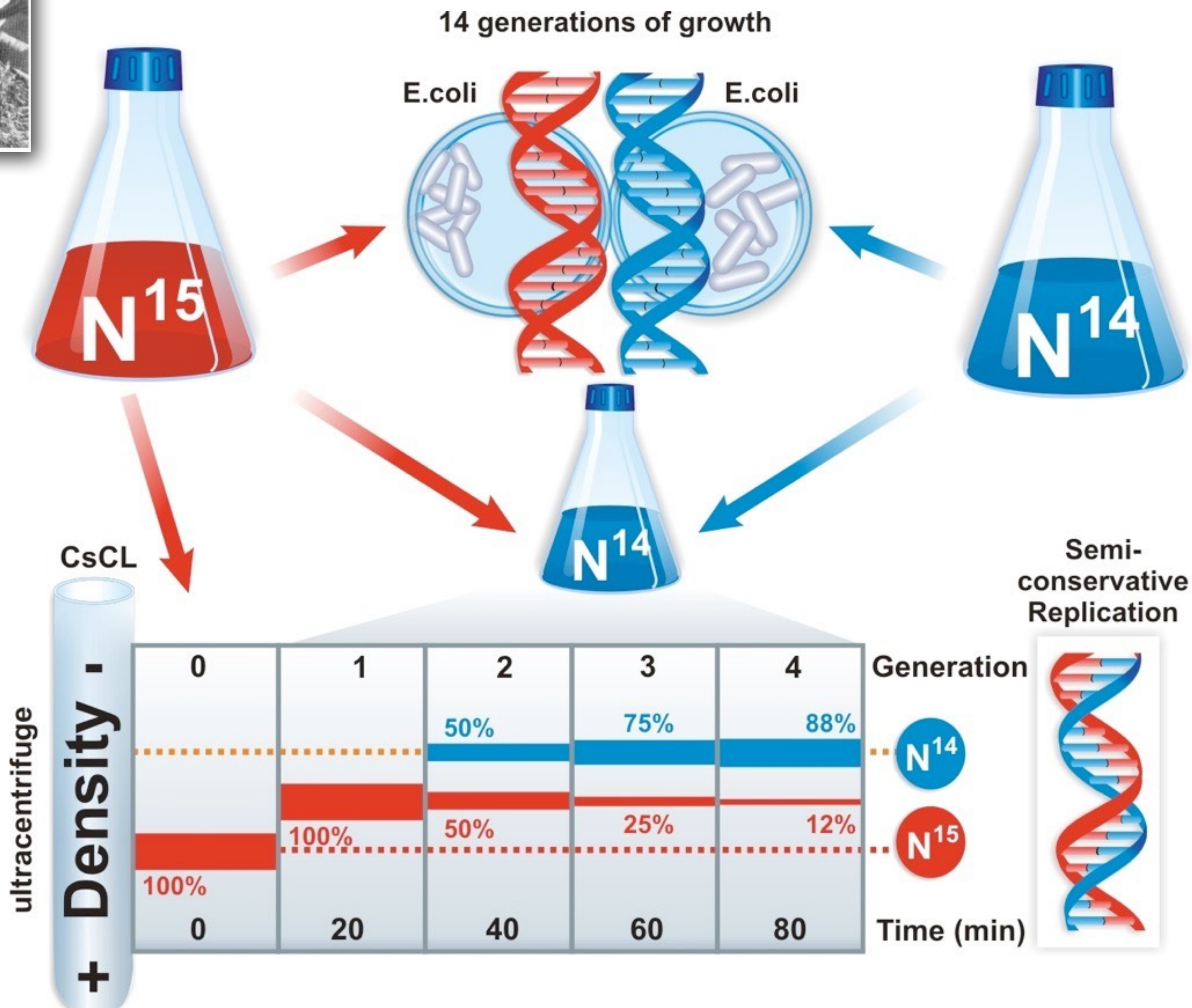
1944 - Oswald Theodore Avery
O DNA carregava a informação genética.

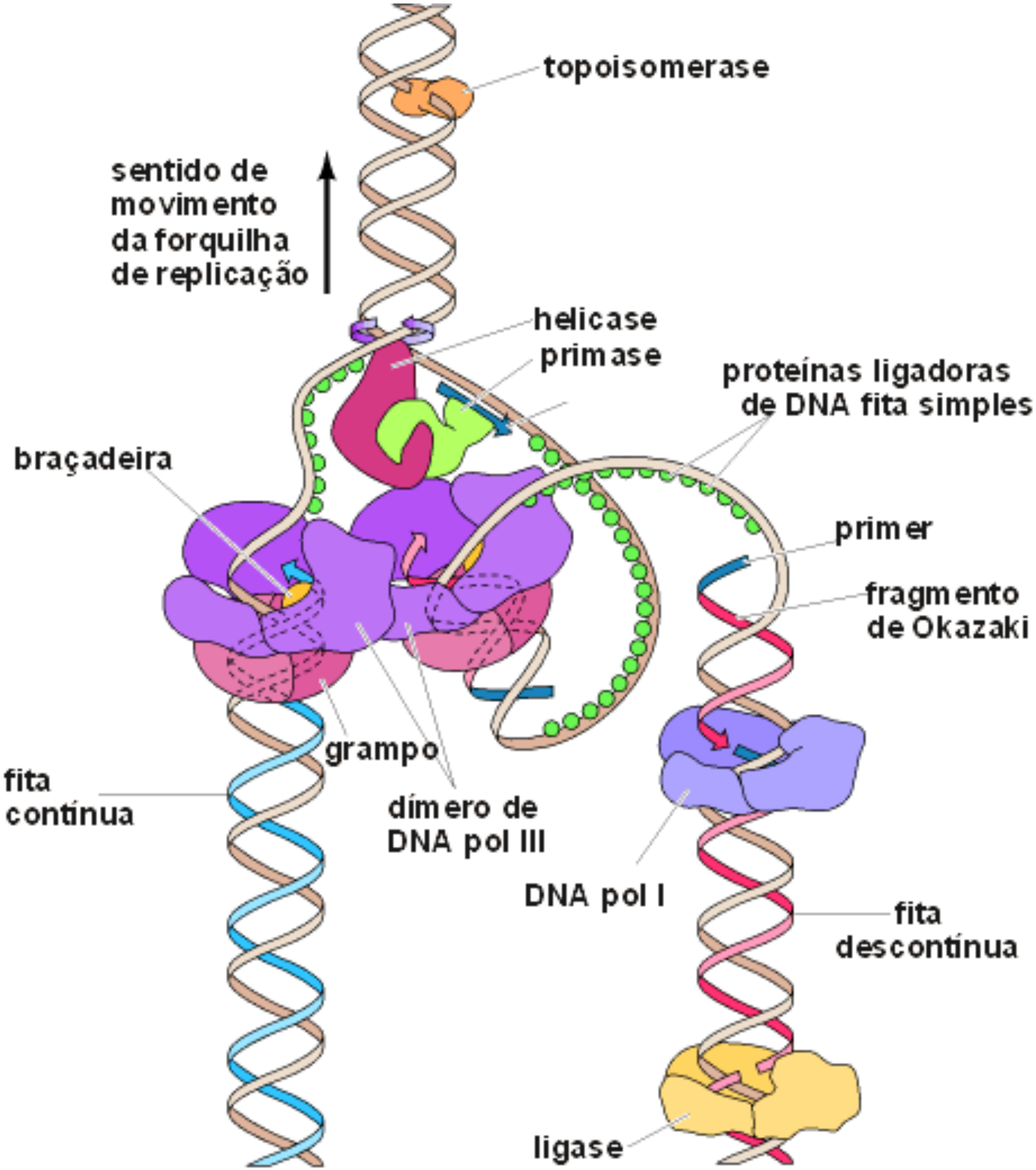


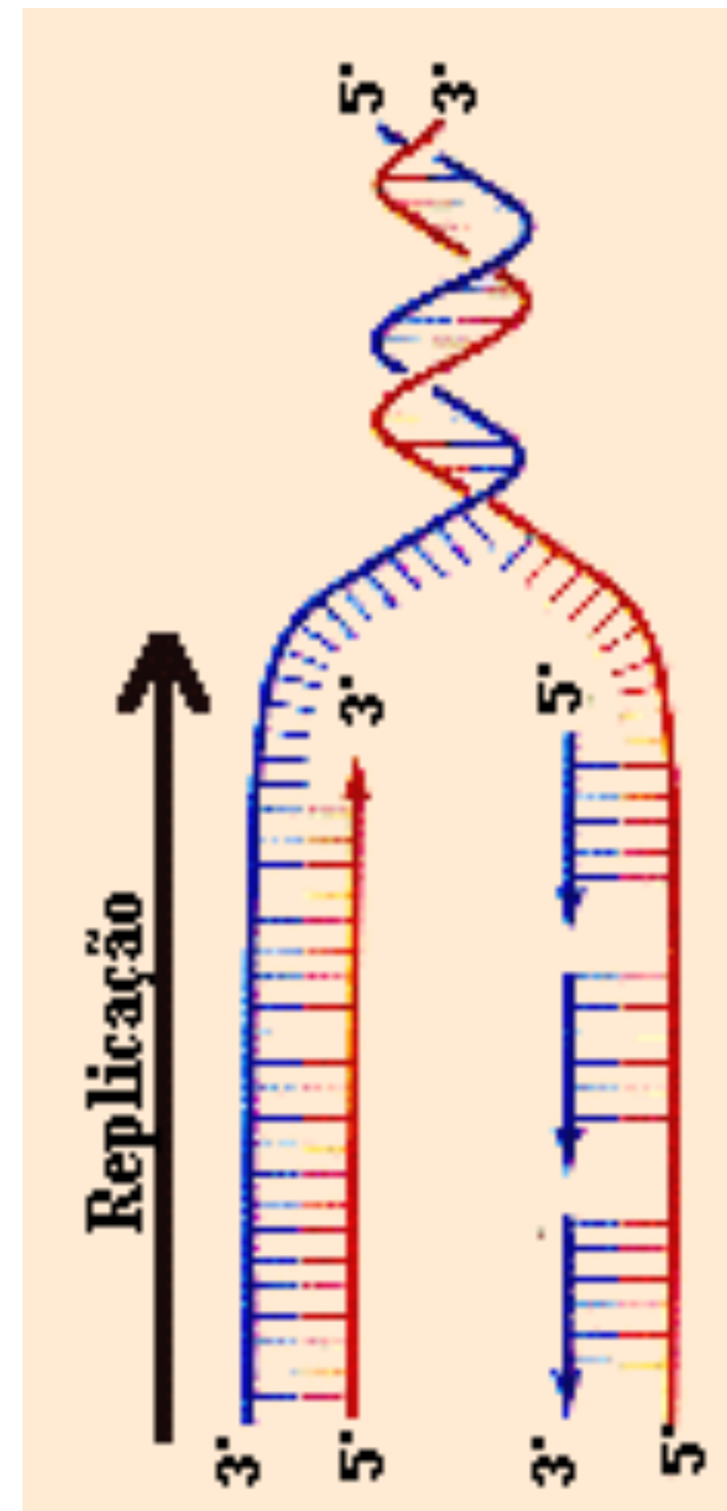
1952 - Hershey-Chase Experiments
Confirmação de que O DNA carregava a informação genética.



1958 - Meselson e Stahl
Modelo de replicação do DNA semiconservativa.




topoisomerase
sentido de movimento da forquilha de replicação
helicase
primase
proteínas ligadoras de DNA fita simples
braçadeira
primer
fragmento de Okazaki
grampo
fita contínua
dímero de DNA pol III
DNA pol I
fita descontínua
ligase

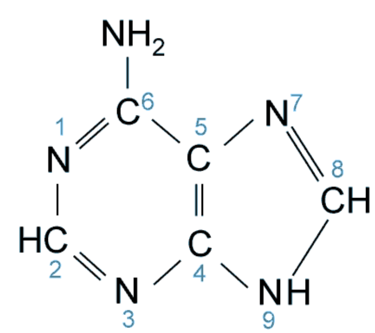

Replicação
5'
3'
3'
5'
3'
5'
3'
5'

DNA

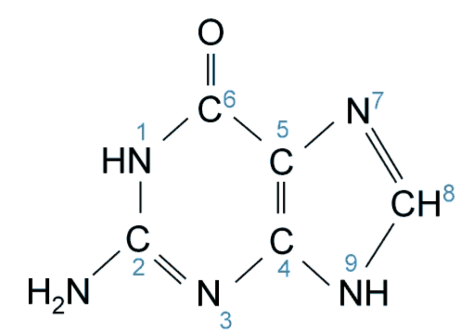
Polímero de nucleotídeos

Bases nitrogenadas

Purinas

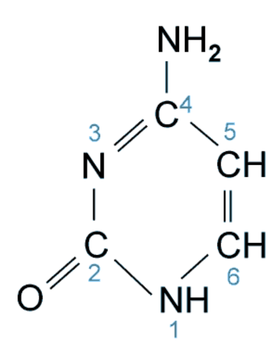


Adenine (A)

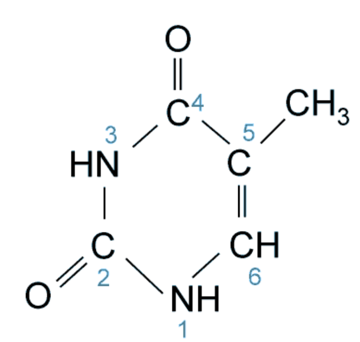


Guanine (G)

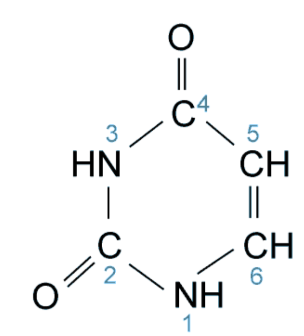
Pirimidinas



Cytosine (C)



Thymine (T)

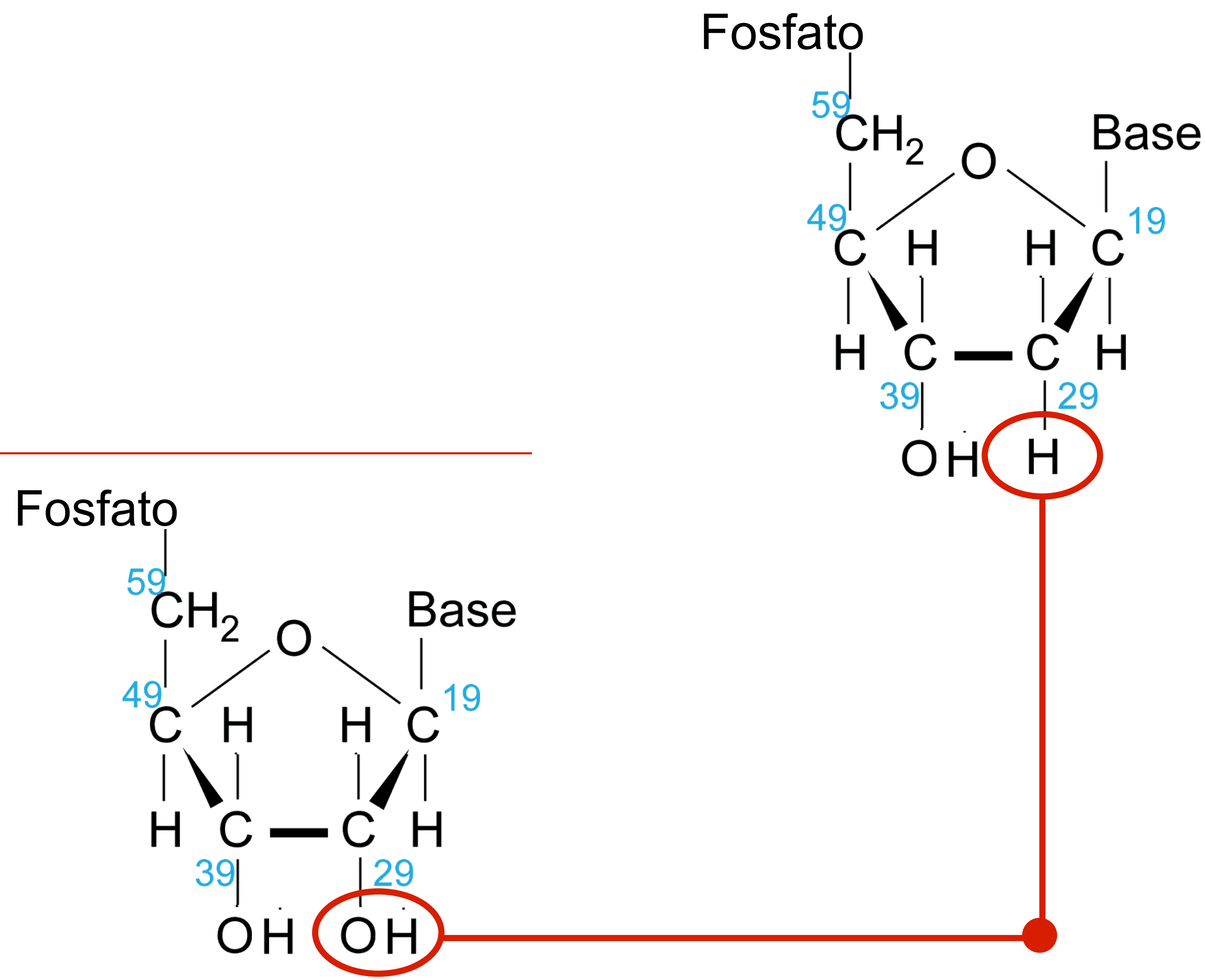


Uracil (U)

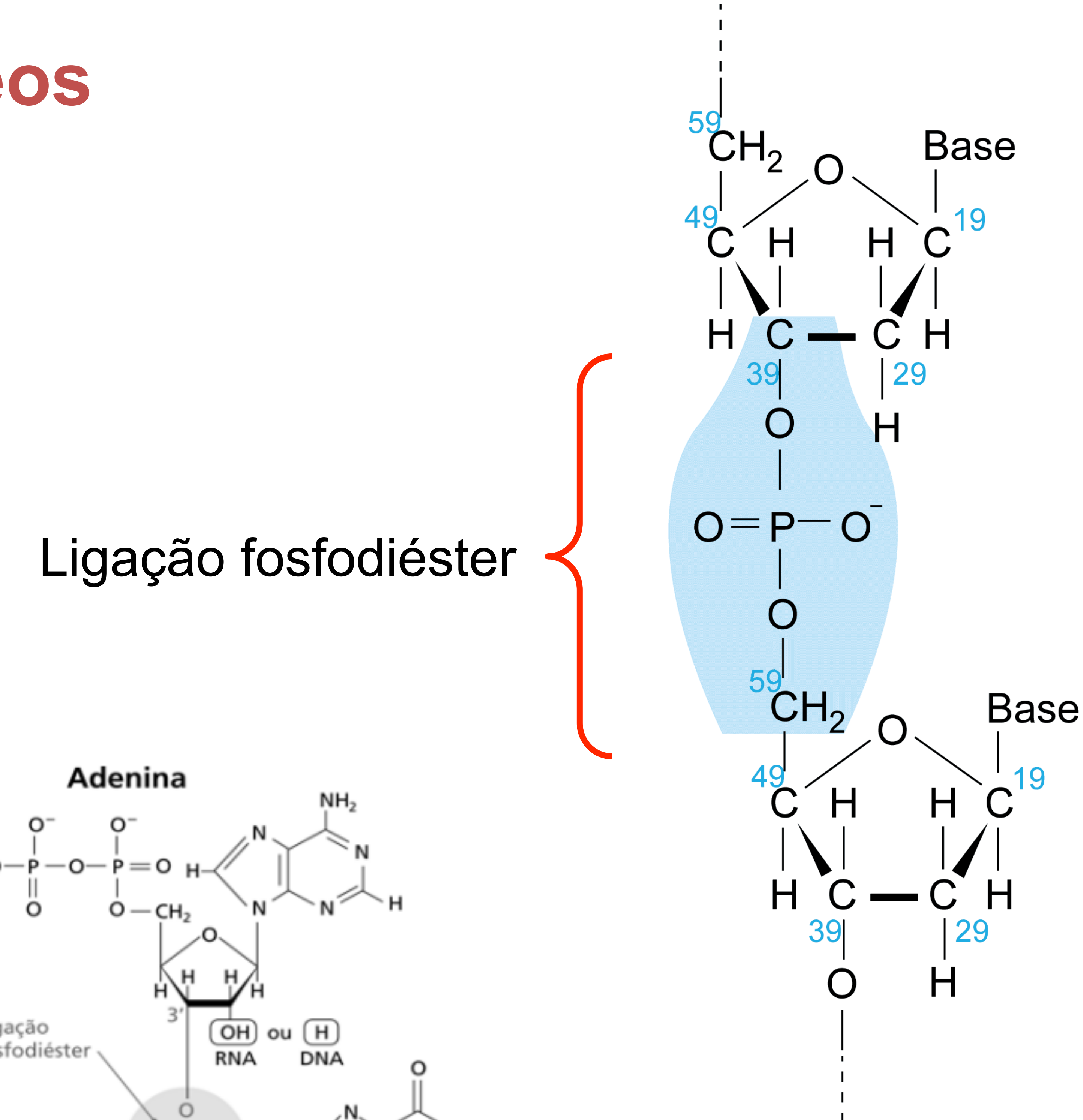
Pentoses

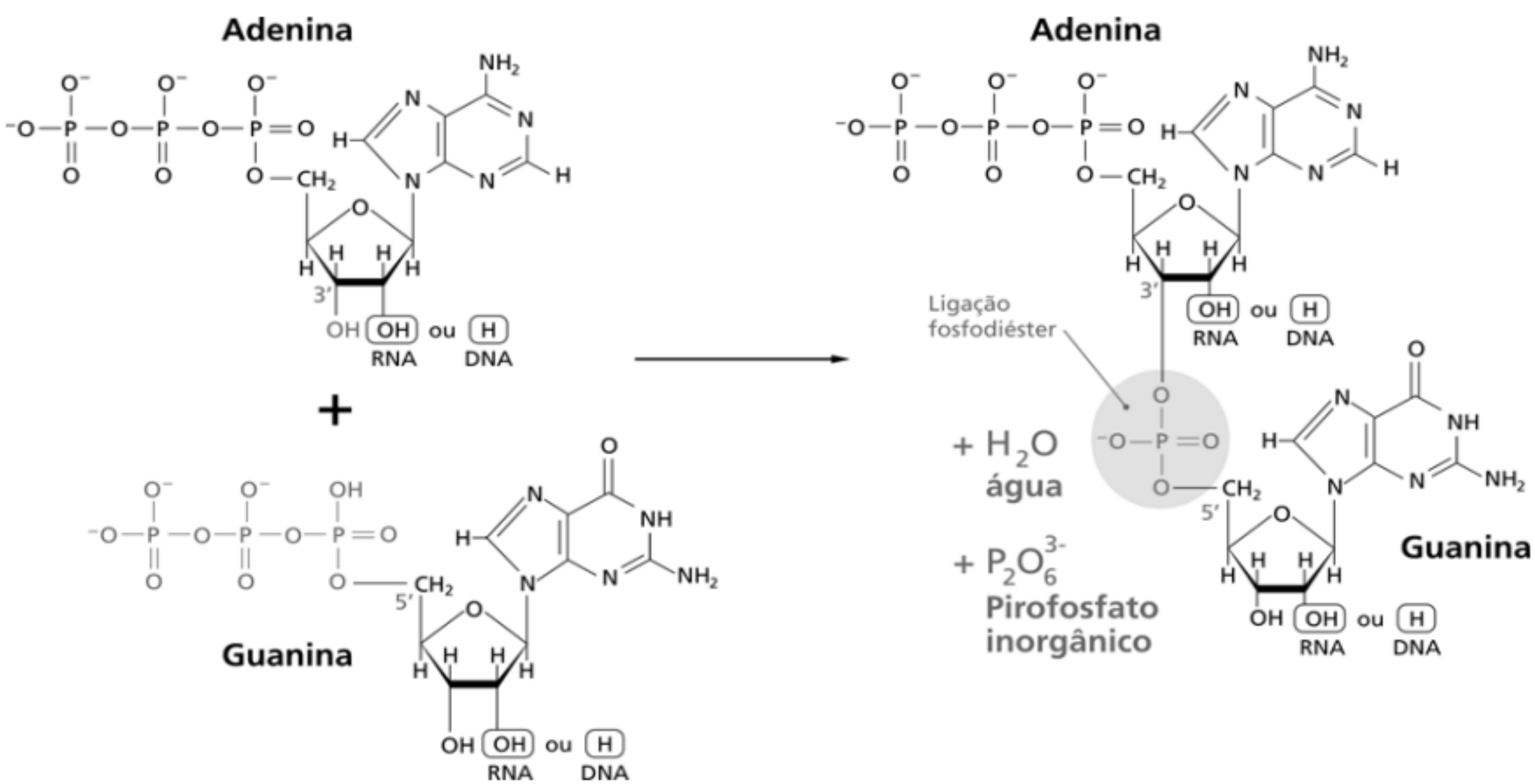
Desoxirribose

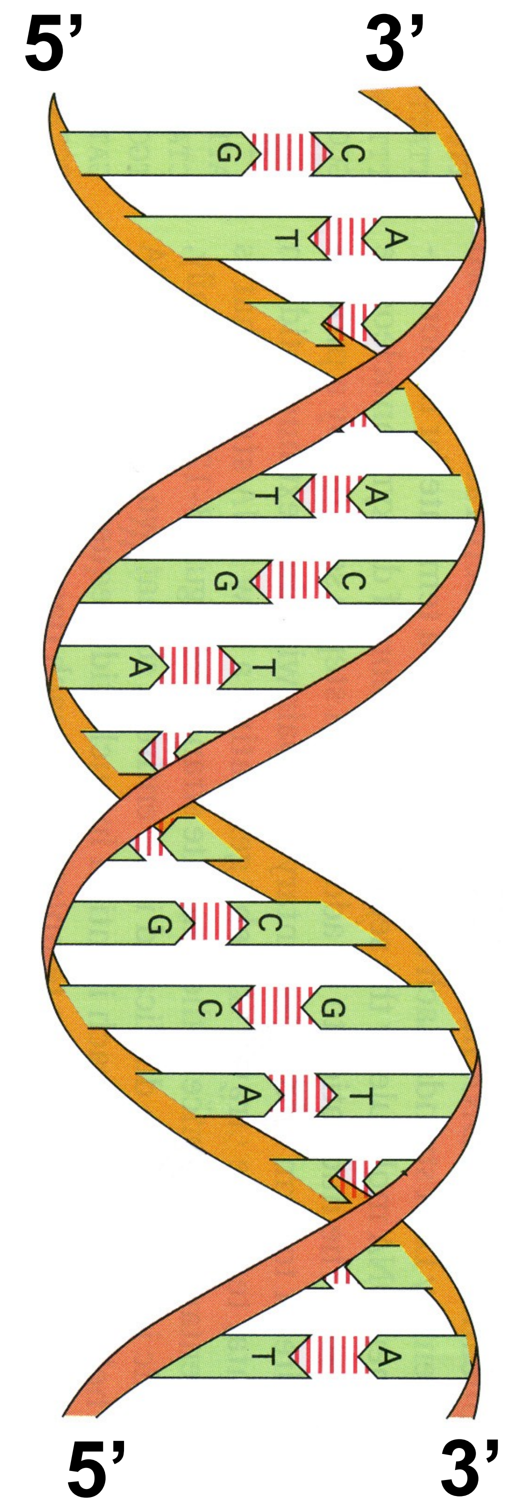
Ribose



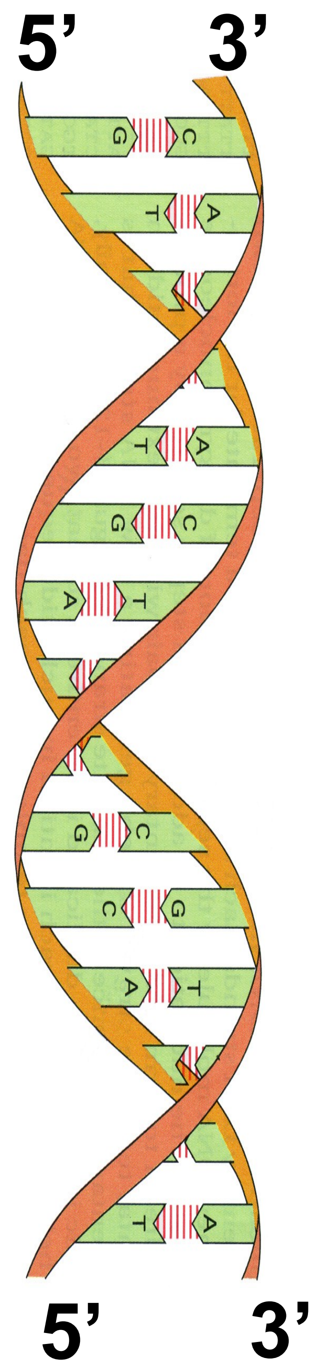
Polinucleotídeos





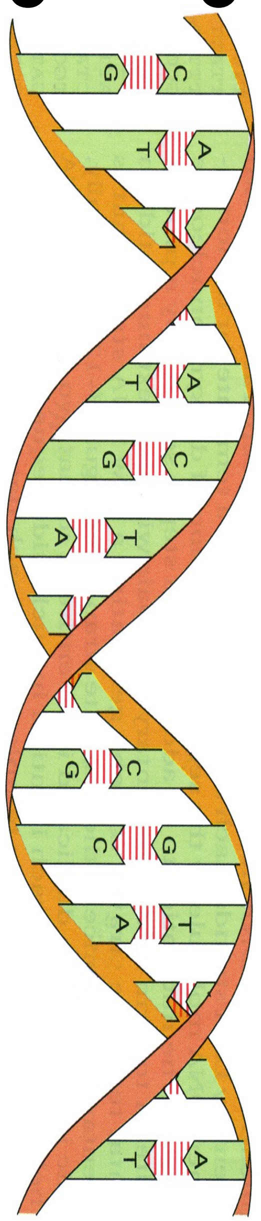


5’
3’
G C
T A
T A
G C
A T
G C
C G
A T
T A
5’
3’



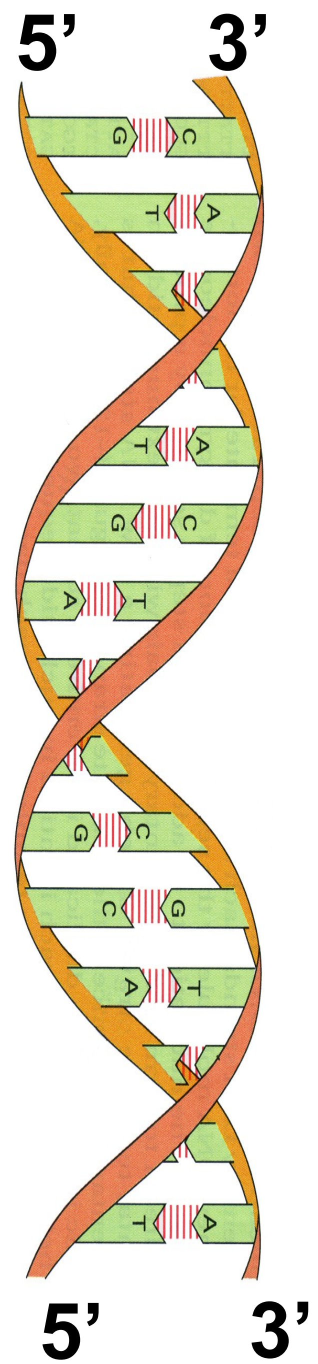
- O DNA contém dois polinucleotídeos;
- As bases nitrogenadas estão posicio-nadas no interior da hélice;
- As bases interagem através de pontes de hidrogênio;
- Cada giro da hélice contém 10 bases;
- Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;
- O DNA possui dois sulcos diferentes;
- O DNA possui giro para a direita.

5’ 3’

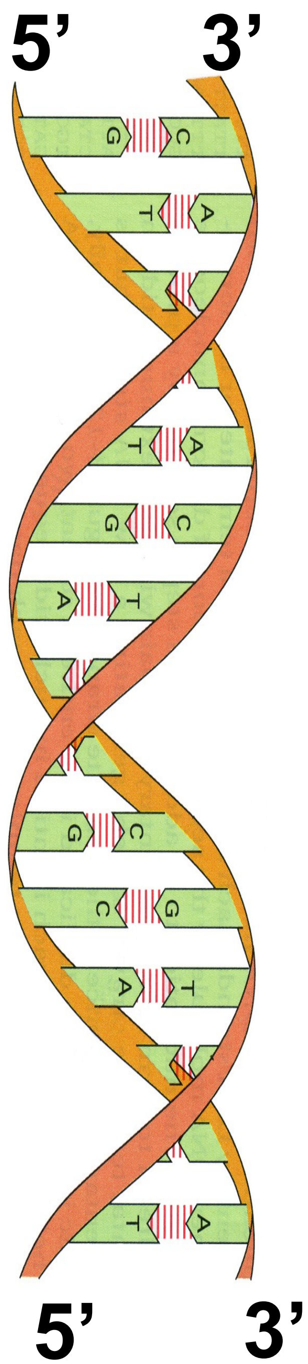


5’ 3’

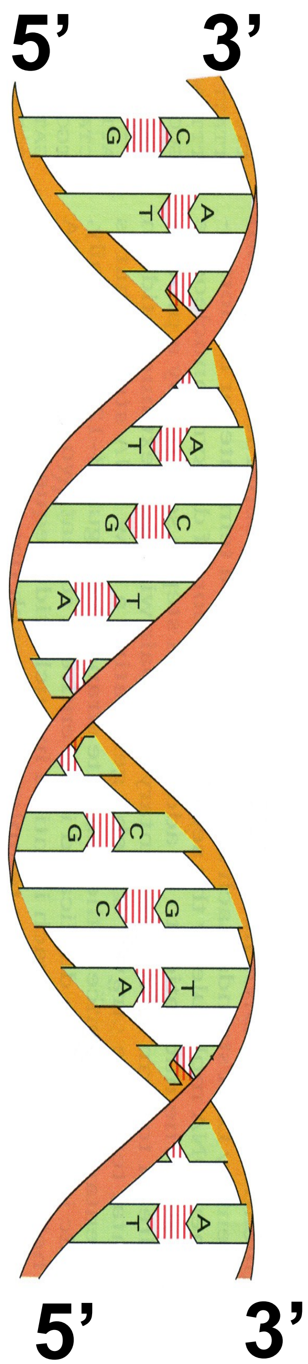
- O DNA contém dois polinucleotídeos;
- **As bases nitrogenadas estão posicionadas no interior da hélice;**
- As bases interagem através de pontes de hidrogênio;
- Cada giro da hélice contém 10 bases;
- Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;
- O DNA possui dois sulcos diferentes;
- O DNA possui giro para a direita.



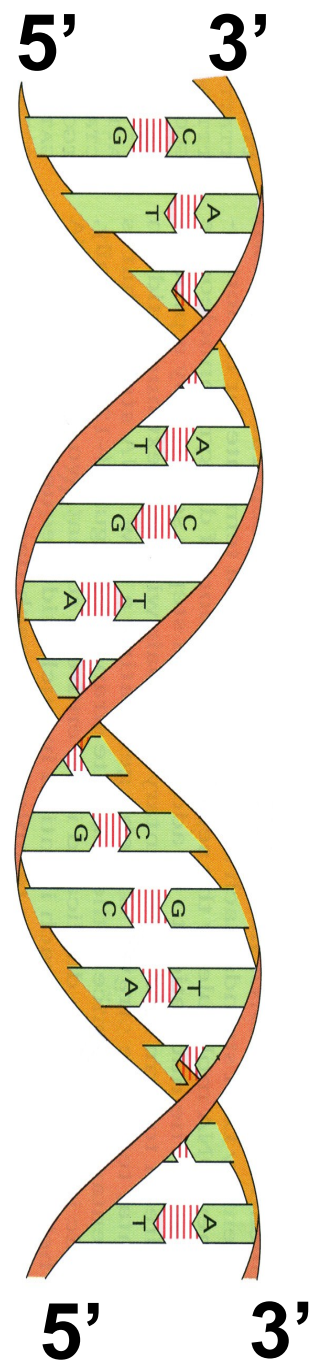
- O DNA contém dois polinucleotídeos;
- As bases nitrogenadas estão posicionadas no interior da hélice;
- **As bases interagem através de pontes de hidrogênio;**
- Cada giro da hélice contém 10 bases;
- Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;
- O DNA possui dois sulcos diferentes;
- O DNA possui giro para a direita.



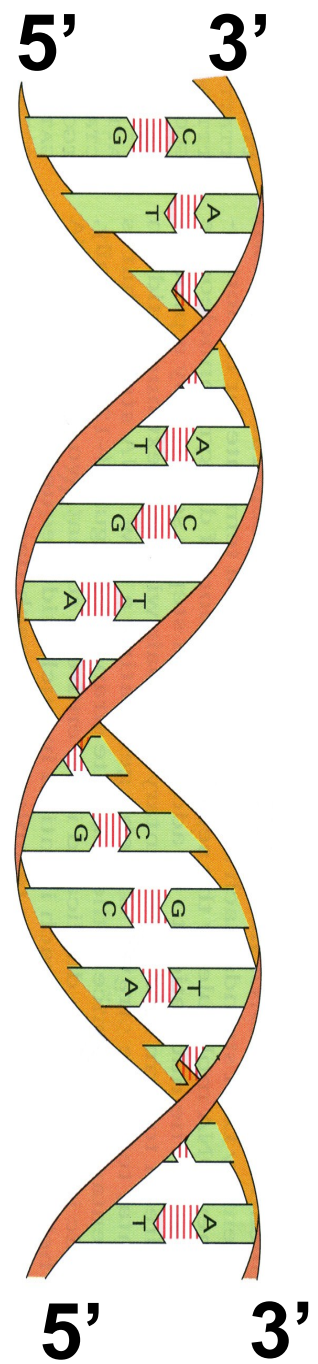
- O DNA contém dois polinucleotídeos;
- As bases nitrogenadas estão posicio-nadas no interior da hélice;
- As bases interagem através de pontes de hidrogênio;
- **Cada giro da hélice contém 10 bases;**
- Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;
- O DNA possui dois sulcos diferentes;
- O DNA possui giro para a direita.



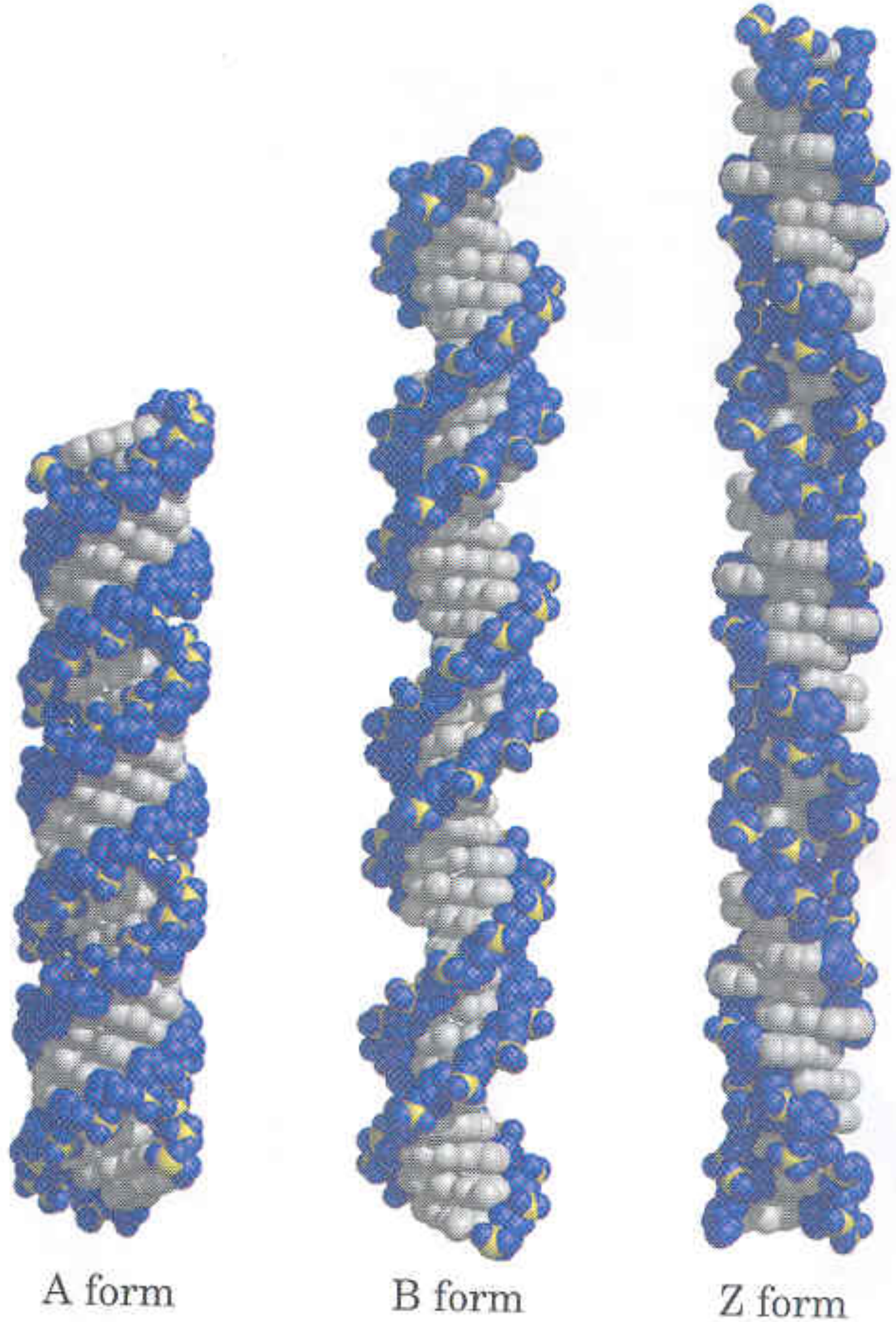
- O DNA contém dois polinucleotídeos;
- As bases nitrogenadas estão posicio-nadas no interior da hélice;
- As bases interagem através de pontes de hidrogênio;
- Cada giro da hélice contém 10 bases;
- **Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;**
- O DNA possui dois sulcos diferentes;
- O DNA possui giro para a direita.



- O DNA contém dois polinucleotídeos;
- As bases nitrogenadas estão posicio-nadas no interior da hélice;
- As bases interagem através de pontes de hidrogênio;
- Cada giro da hélice contém 10 bases;
- Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;
- **O DNA possui dois sulcos diferentes;**
- O DNA possui giro para a direita.



- O DNA contém dois polinucleotídeos;
- As bases nitrogenadas estão posicio-nadas no interior da hélice;
- As bases interagem através de pontes de hidrogênio;
- Cada giro da hélice contém 10 bases;
- Os dois polinucleotídeos têm sentidos inversos ou anti-paralelo;
- O DNA possui dois sulcos diferentes;
- **O DNA possui giro para a direita.**


A form
B form
Z form

Wener Arber
Enzimas de restrição

Martin Gellert
DNA ligase

Boyer, Cohen e Berg
Clonagem gênica

1962 1966 1967 1972

Niremberg, Ochoa e Khorana
Elucidaram o código genético

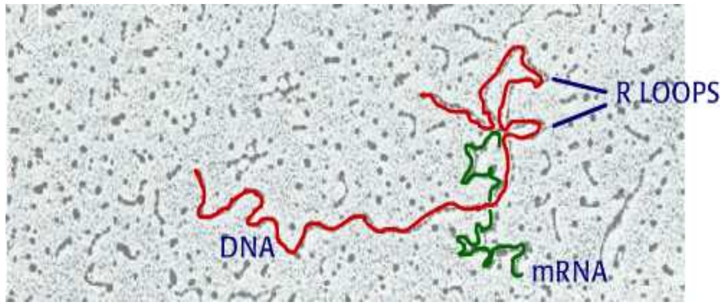
	T			C			A			G		
T	TTT	Phe	F	TCT	Ser	S	TAT	Tyr	Y	TGT	Cys	C
	TTC			TCC			TAC			TGC		
	TTA	Leu	L	TCA			TAA	STOP		TGA	STOP	
	TTG			TCG			TAG			TGG	Trp	W
C	CTT	Leu	L	CCT	Pro	P	CAT	His	H	CGT	Arg	R
	CTC			CCC			CAC			CGC		
	CTA			CCA			CAA	Gln	Q	CGA		
	CTG			CCG			CAG			CGG		
A	ATT	Ile	I	ACT	Thr	T	AAT	Asn	N	AGT	Ser	S
	ATC			ACC			AAC			AGC		
	ATA			ACA			AAA	Lys	K	AGA	Arg	R
	ATG	Met	M	ACG			AAG			AGG		
G	GTT	Val	V	GCT	Ala	A	GAT	Asp	D	GGT	Gly	G
	GTC			GCC			GAC			GGC		
	GTA			GCA			GAA	Glu	E	GGA		
	GTG			GCG			GAG			GGG		

	T			C			A			G		
T	TTT	Phe	F	TCT	Ser	S	TAT	Tyr	Y	TGT	Cys	C
	TTC			TCC			TAC			TGC		
	TTA	Leu	L	TCA			TAA	STOP		TGA	STOP	
	TTG			TCG			TAG			TGG	Trp	W
C	CTT	Leu	L	CCT	Pro	P	CAT	His	H	CGT	Arg	R
	CTC			CCC			CAC			CGC		
	CTA			CCA			CAA	Gln	Q	CGA		
	CTG			CCG			CAG			CGG		
A	ATT	Ile	I	ACT	Thr	T	AAT	Asn	N	AGT	Ser	S
	ATC			ACC			AAC			AGC		
	ATA			ACA			AAA	Lys	K			
	ATG	Met	M	ACG			AAG					
G	GTT	Val	V	GCT	Ala	A	GAT	Asp	D			
	GTC			GCC			GAC					
	GTA			GCA			GAA	Glu	E			
	GTG			GCG			GAG					

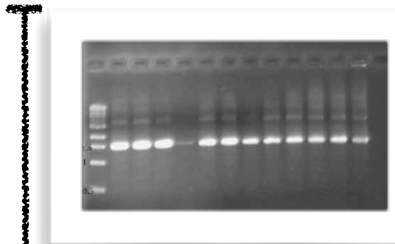
Codon Usage in Homo sapiens

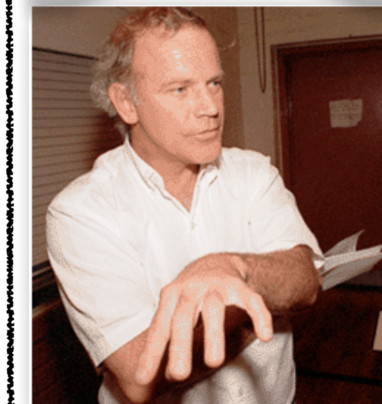
UUU F 0.46	UCU S 0.19	UAU Y 0.44	UGU C 0.46
UUC F 0.54	UCC S 0.22	UAC Y 0.56	UGC C 0.54
UUA L 0.08	UCA S 0.15	UAA * 0.30	UGA * 0.47
UUG L 0.13	UCG S 0.05	UAG * 0.24	UGG W 1.00
CUU L 0.13	CCU P 0.29	CAU H 0.42	CGU R 0.08
CUC L 0.20	CCC P 0.32	CAC H 0.58	CGC R 0.18
CUA L 0.07	CCA P 0.28	CAA Q 0.27	CGA R 0.11
CUG L 0.40	CCG P 0.11	CAG Q 0.73	CGG R 0.20
AUU I 0.36	ACU T 0.25	AAU N 0.47	AGU S 0.15
AUC I 0.47	ACC T 0.36	AAC N 0.53	AGC S 0.24
AUA I 0.17	ACA T 0.28	AAA K 0.43	AGA R 0.21
AUG M 1.00	ACG T 0.11	AAG K 0.57	AGG R 0.21
GUU V 0.18	GCU A 0.27	GAU D 0.46	GGU G 0.16
GUC V 0.24	GCC A 0.40	GAC D 0.54	GGC G 0.34
GUA V 0.12	GCA A 0.23	GAA E 0.42	GGA G 0.25
GUG V 0.46	GCG A 0.11	GAG E 0.58	GGG G 0.25

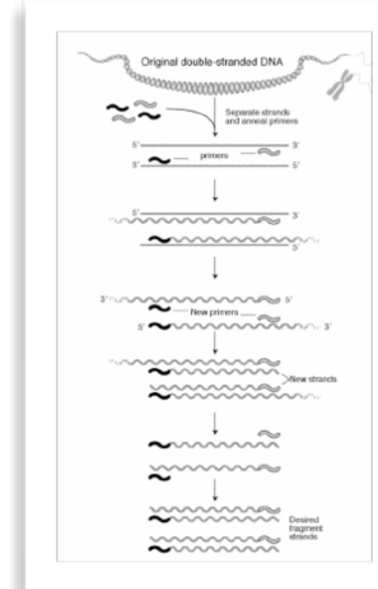
[Codon/a.a./fraction per codon per a.a.]
Homo sapiens data from the Codon Usage Database



Phillip Sharp
Genes de eucariotos são interrompidos

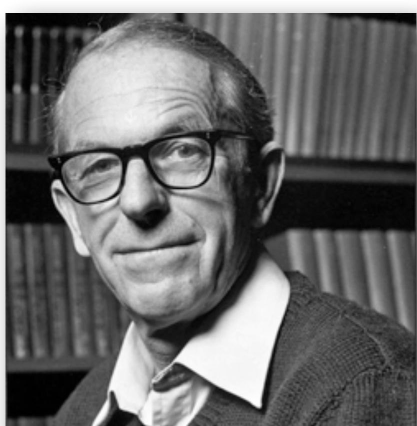


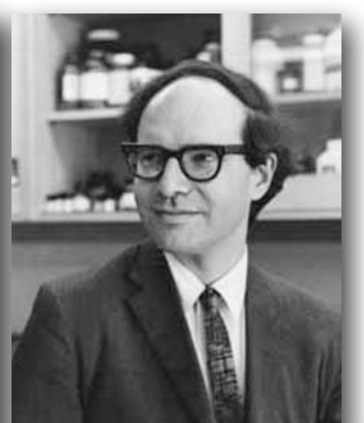




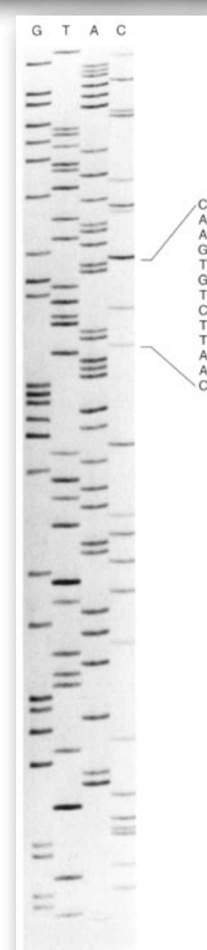
Kary Mullis
Reação em cadeia da polimerase (PCR)

1975-1977 1984 1986 1987





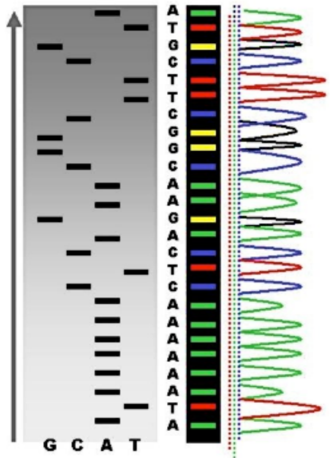
Frederick Sanger e Walter Gilbert
Sequenciamento do DNA



1986
The Human Genome Initiative, later called the Human Genome Project, is announced. The goal is to sequence the entire human genome and provide a complete catalog of every human gene.

1984
Alec Jeffreys develops "genetic fingerprinting" a molecular biological analog of traditional fingerprinting for identifying individuals by analyzing polymorphic sequences in the DNA.

Leroy Hood



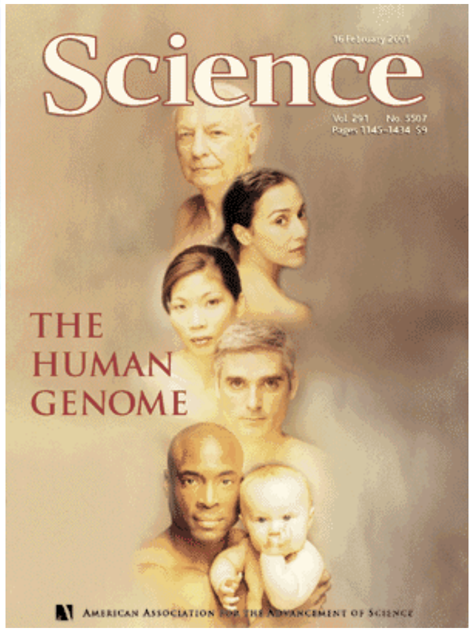
Francis Collins

Diretor do "National Human Genome Research Institute"

1990 1999 2001

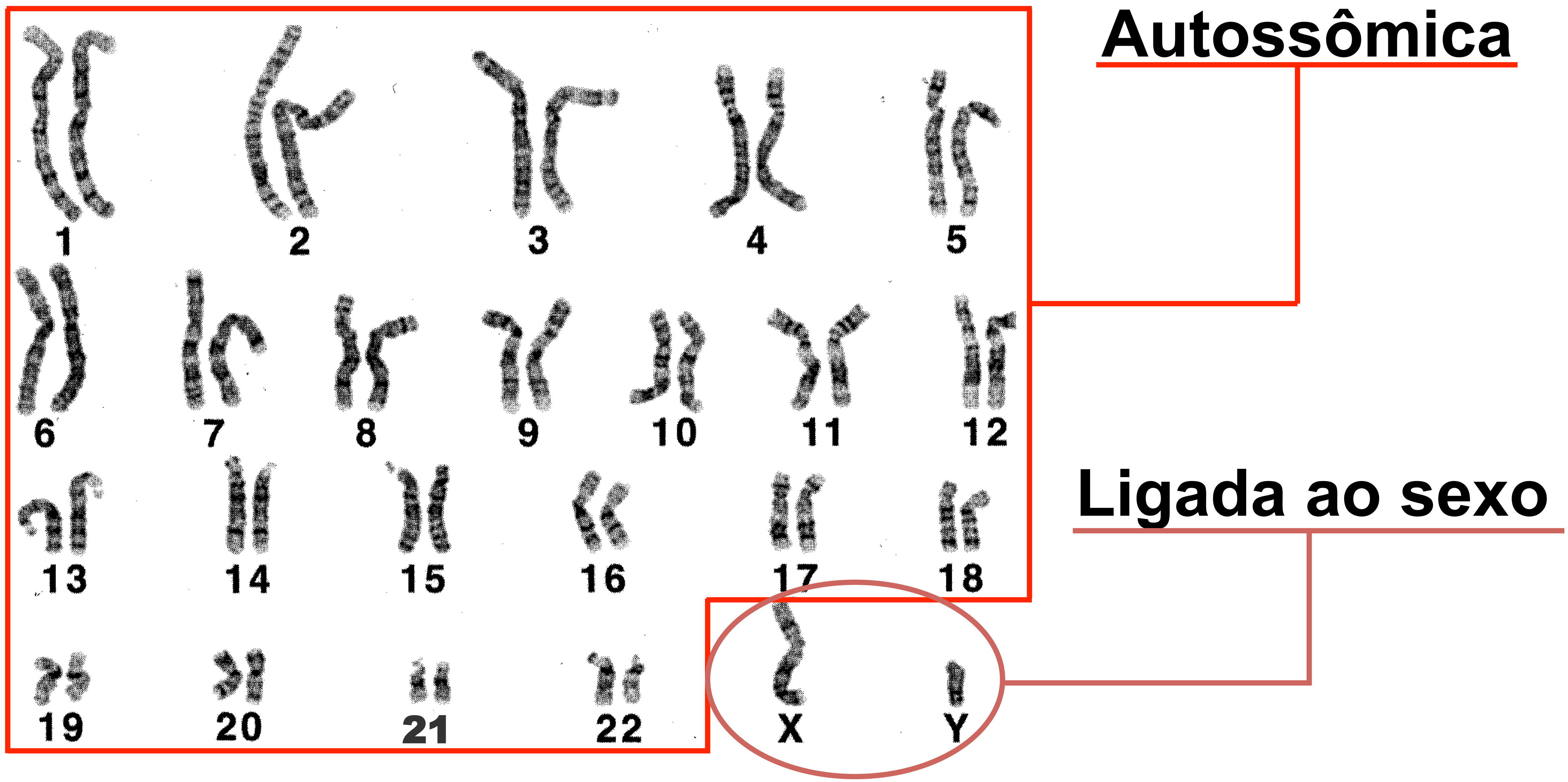
Craig Venter

Co-fundador da "Celera Genomics Corporation"



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.


Autossômica
Ligada ao sexo
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y