

Biology I
Basic Principles of Cell Biology

DNA as the Carrier of Genetic Information

The structure and role of DNA

Yiannis Drosos
Assistant Professor
Laboratory of Biology
Department of Medicine
University of Patras
idosos@upatras.gr



One letter in 3 billion

β -globin gene (HBB) · coding strand, 5'→3'

Normal allele (HbA)

ATG	GTG	CAC	CTG	ACT	CCT	GAG	GAG	AAG
Met	Val	His	Leu	Thr	Pro	Glu	Glu	Lys

Sickle cell allele (HbS)

ATG	GTG	CAC	CTG	ACT	CCT	GTG	GAG	AAG
Met	Val	His	Leu	Thr	Pro	Val	Glu	Lys

1 base pair → GAG→GTG (Glu6Val) → HbS polymerization → sickled red cells, vaso-occlusive crises

How can one molecule store, copy and pass on information with such precision?

2023

First CRISPR-Cas9 therapy (exa-cel) approved for sickle cell disease and β -thalassemia (UK: November; USA: December 2023)

2025

First personalized in vivo base-editing therapy, given to an infant with CPS1 deficiency and designed for his own mutation

Three questions for today



1

How do we know that DNA is the genetic material?

→ The key experiments, 1928–1952

2

How does the structure of DNA explain its function?

→ Nucleotides, base pairing, double helix

3

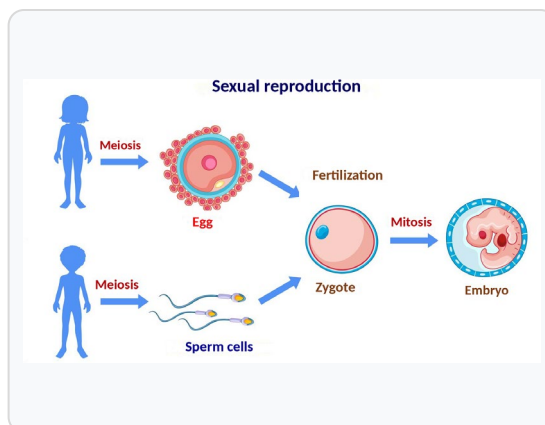
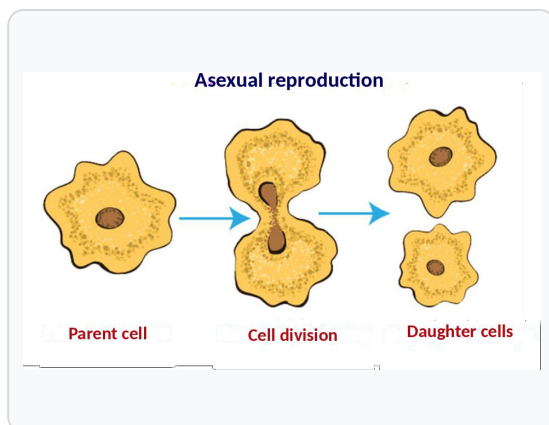
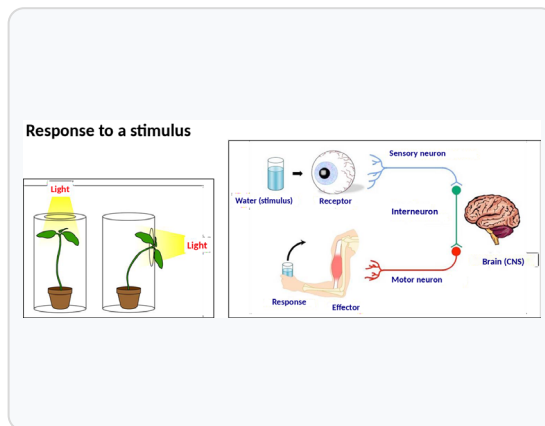
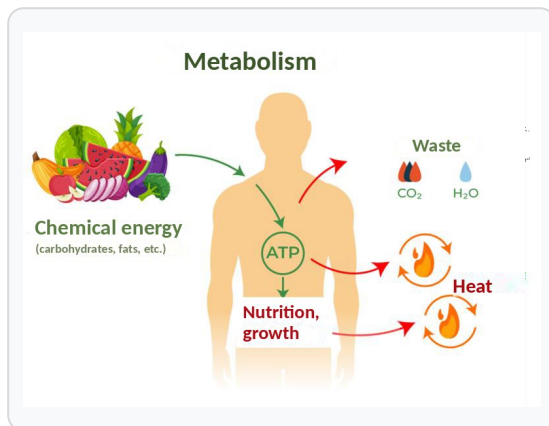
How is genetic information “written”?

→ Sequence, gene, genome



Next lecture: how DNA is packaged into chromatin and chromosomes · Essential Cell Biology, Ch. 5

What must a genetic material do?



1

Storage

Stably encodes a huge amount of information

2

Replication

Is copied faithfully at every cell division

3

Expression

Is “read” to make RNA and proteins

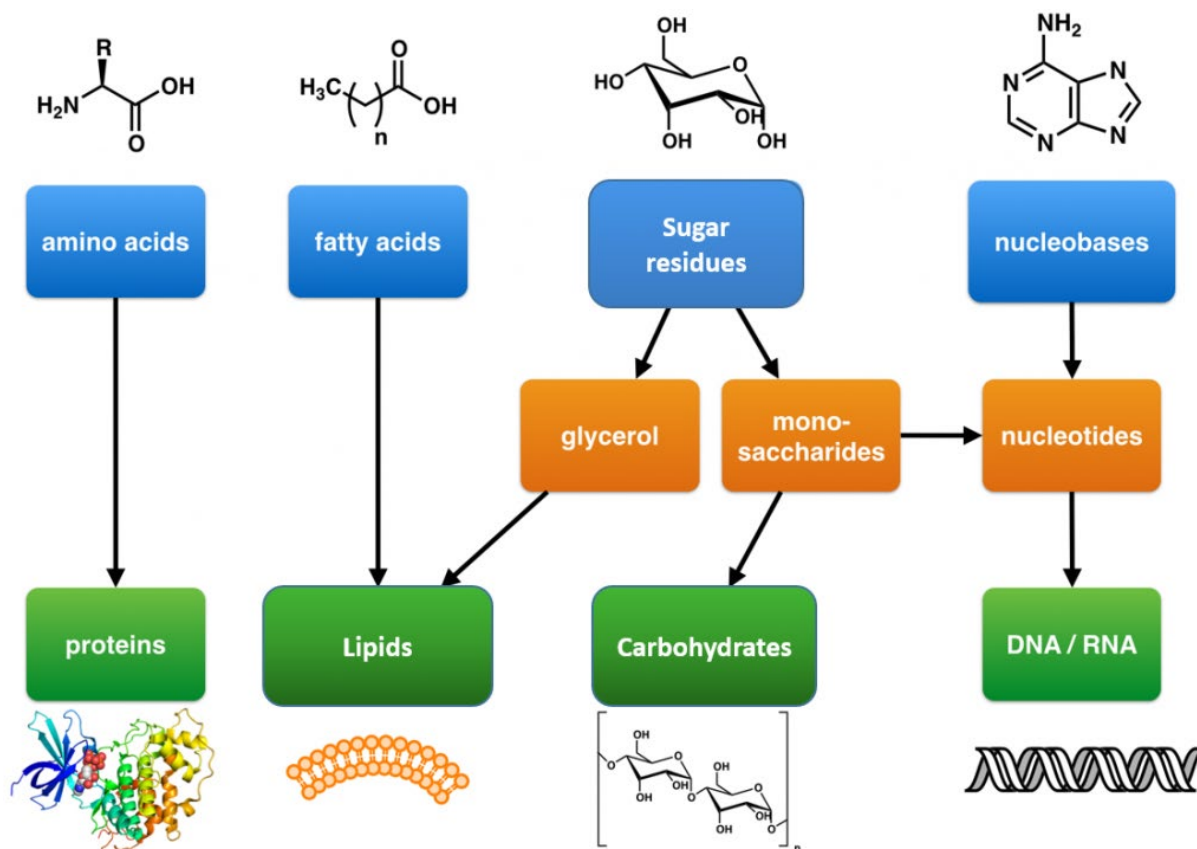
4

Variation

Allows rare changes (mutations) → evolution

Basic properties of life: metabolism, response to stimuli, reproduction

1940: proteins or DNA?



From monomers to the polymers of the cell (Ch. 2)

20

Proteins

20 amino acids, huge chemical diversity → the "obvious" candidate

4

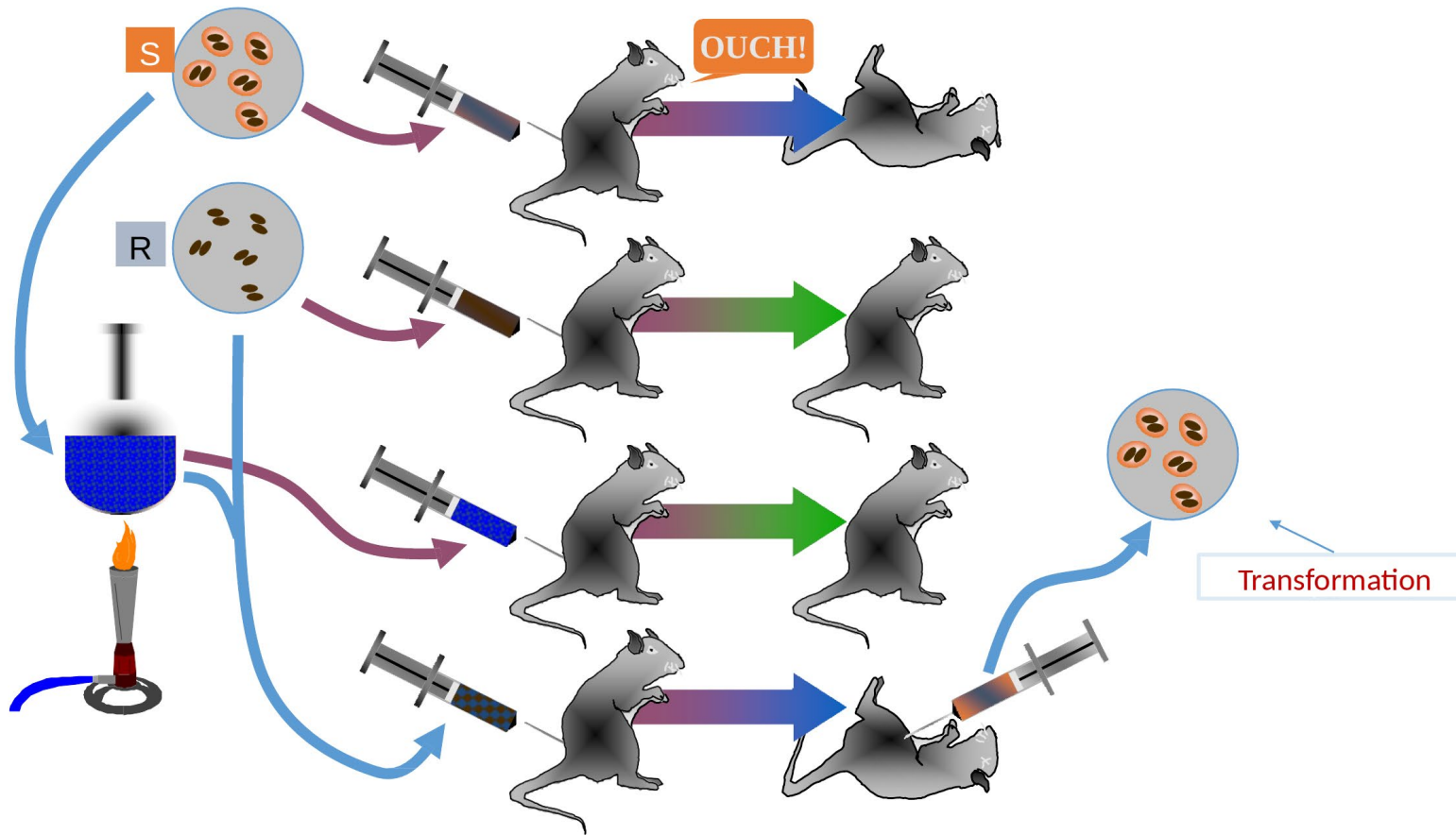
DNA

Only 4 nucleotides; thought to be monotonous (the "tetranucleotide hypothesis", P. Levene)

1869: F. Miescher isolates "nuclein" from white blood cell nuclei

The answer came from experiments, not from intuition.

Griffith (1928): bacterial transformation

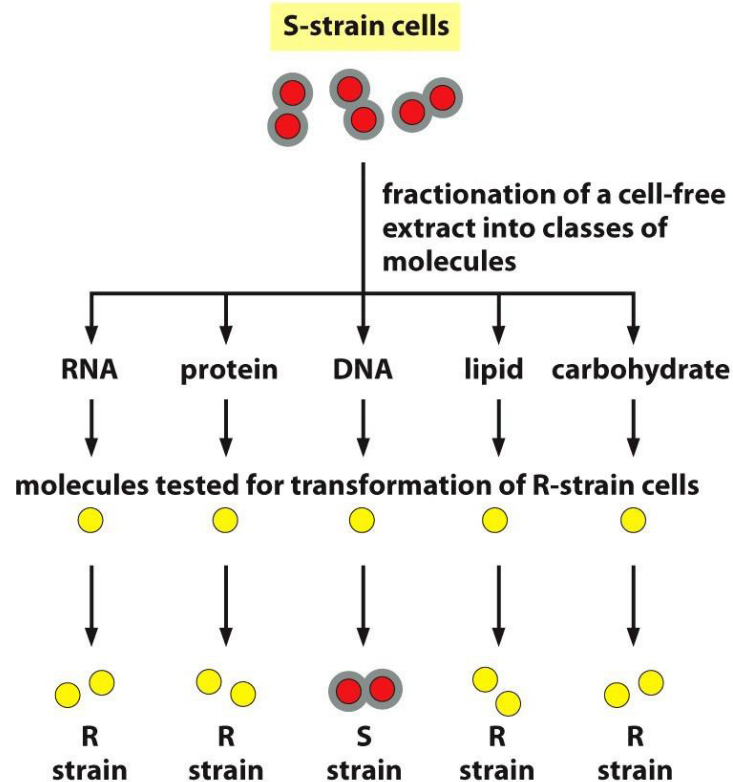


- **S (smooth):** polysaccharide capsule → virulent
- **R (rough):** no capsule → harmless
- Heat-killed S + live R → **death**; live S in the blood
- The R → S change is permanent: it is inherited

“Transforming principle”

Something heritable passes from the dead S cells to the R cells. But which molecule?

Avery, MacLeod & McCarty (1944): it is DNA



CONCLUSION: The molecule that carries the heritable information is DNA.



O. Avery · C. MacLeod · M. McCarty

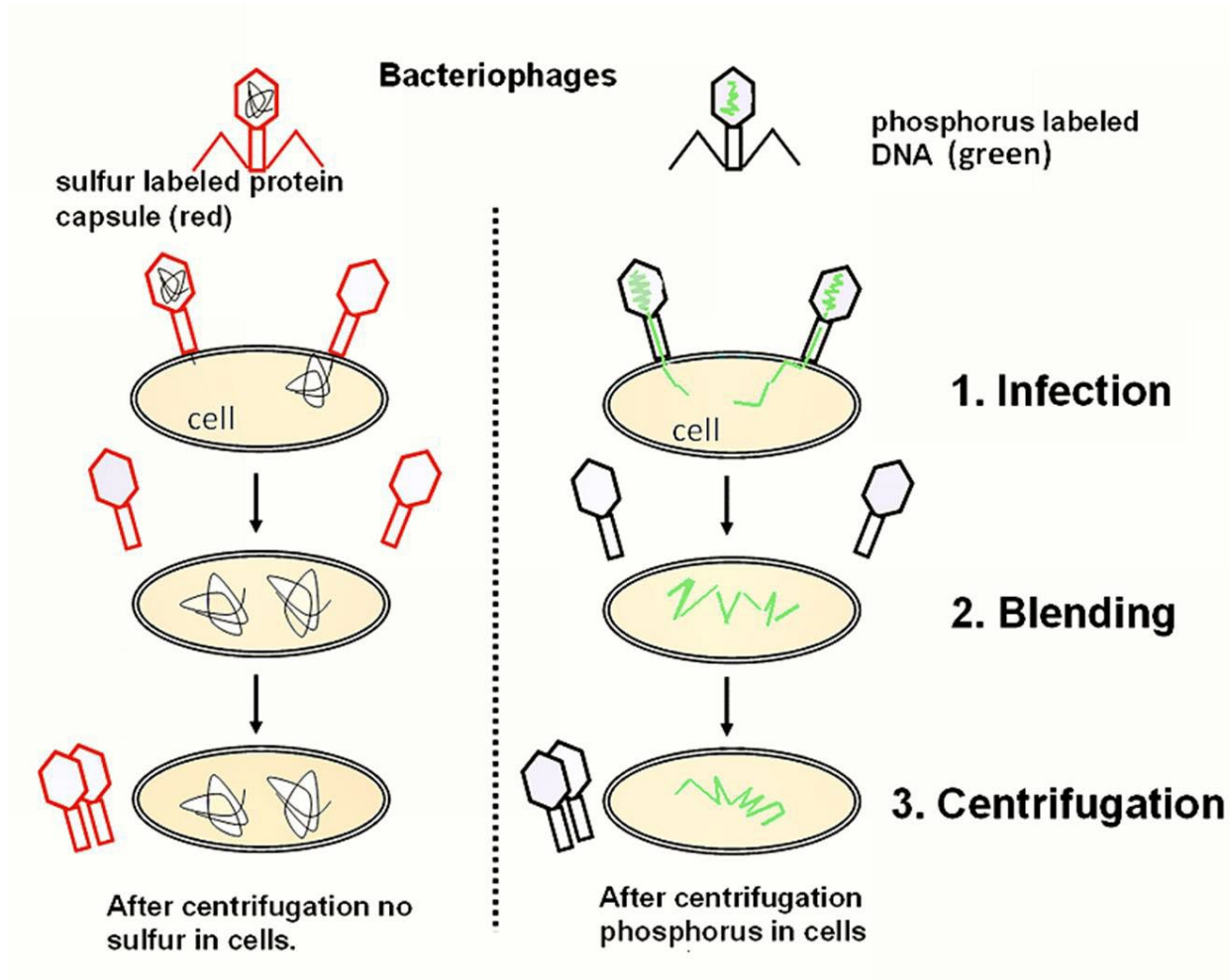
Fractionation of the S-cell extract: only the DNA fraction transforms R into S.

Treatment of active fraction	Transformation?
Proteases	Yes
Ribonuclease (RNase)	Yes
Deoxyribonuclease (DNase)	No

Conclusion

DNA permanently converts R into S: it is the carrier of heritable information.

Hershey & Chase (1952): DNA is inherited



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



A. Hershey & M. Chase

Phage T2: only protein + DNA

$^{35}\text{S} \rightarrow \text{protein}$ · $^{32}\text{P} \rightarrow \text{DNA}$

75% vs 15%

removed by blending (^{35}S vs ^{32}P):
the protein stays outside

<1% vs ≥30%

passed to progeny phage (^{35}S vs ^{32}P)

DNA, not protein, enters the cell and is inherited.

Make a prediction

Suppose the genetic material of phage T2 were protein. What would we find in the progeny phage?



^{32}P but not ^{35}S



^{35}S but not ^{32}P



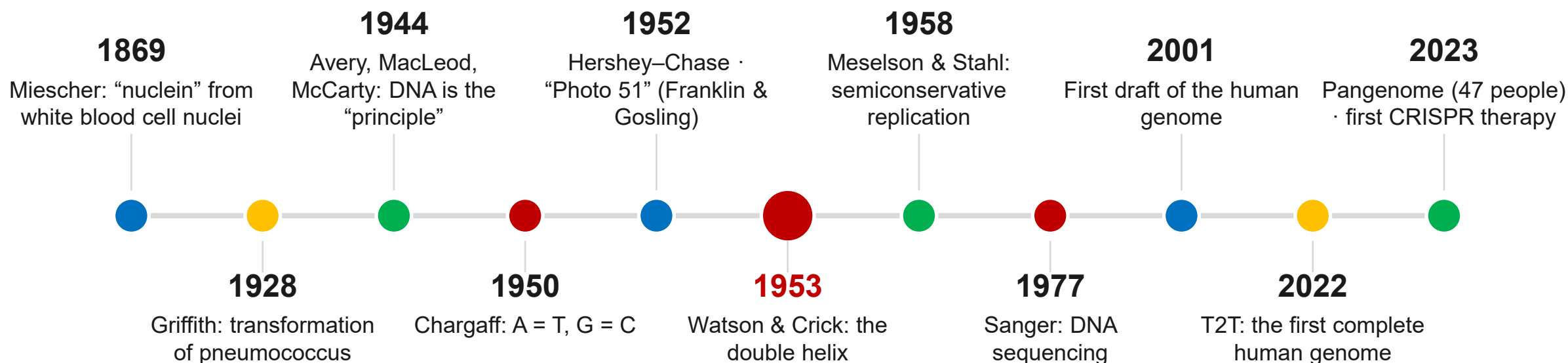
Neither ^{35}S nor ^{32}P



^{35}S and ^{32}P in equal amounts



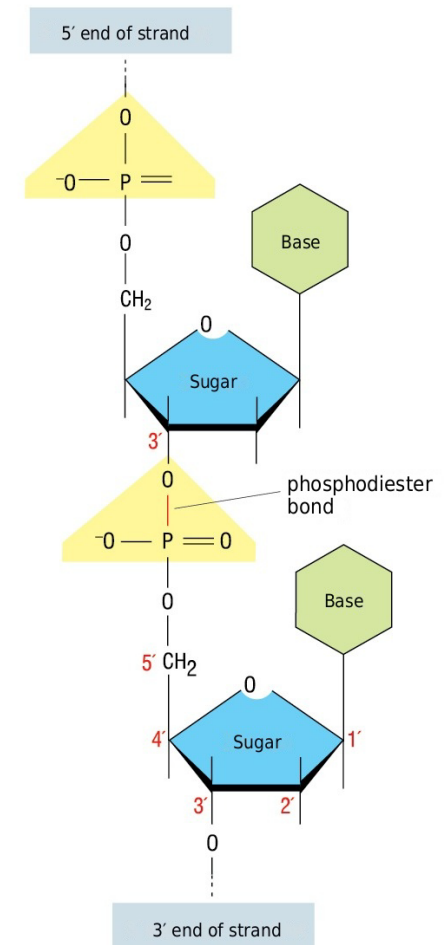
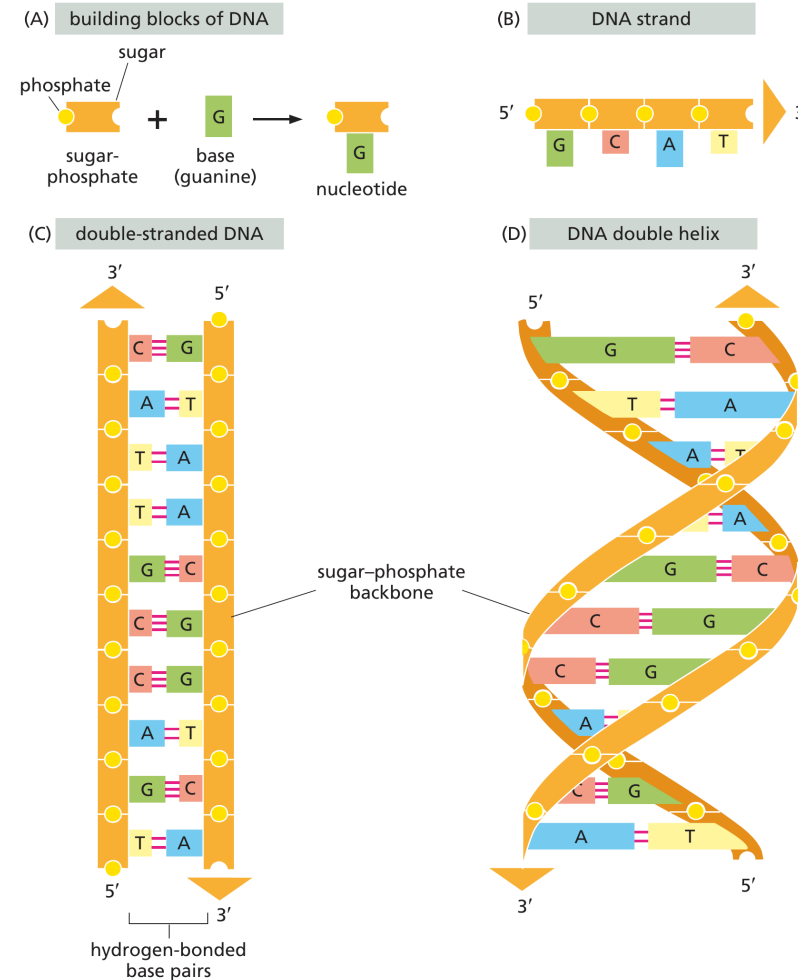
From “nuclein” to the complete genome



The axis is not to scale!

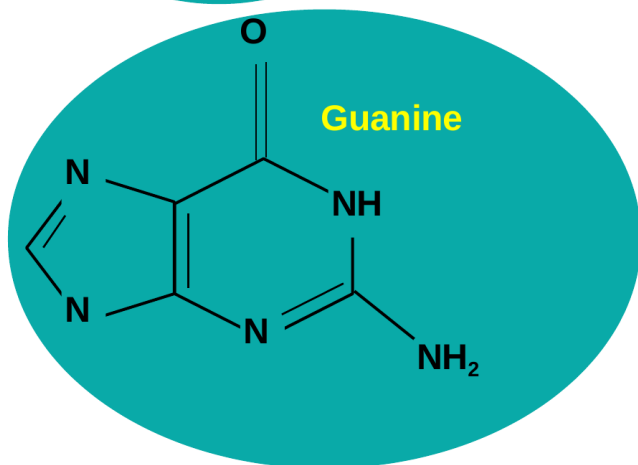
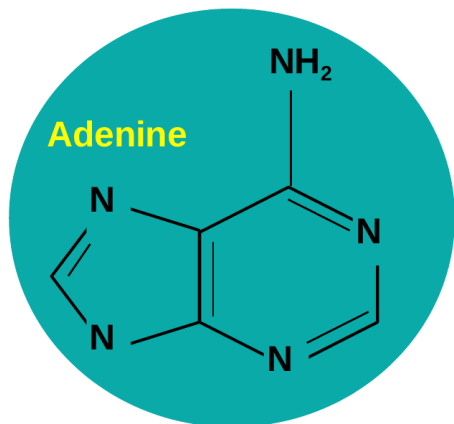
The building blocks: nucleotides

- **Nucleotide** = phosphate group + pentose sugar + nitrogenous base
- **Nucleoside** = sugar + base (no phosphate)
- **Sugar carbons 1'–5'**: base on C1', phosphate on C5'
- **2'-Deoxyribose** (DNA: 2'-H) vs ribose (RNA: 2'-OH) → DNA is chemically more stable
- **Phosphodiester bond**: 3'-OH → phosphate → 5' of the next ⇒ **5'→3' polarity**

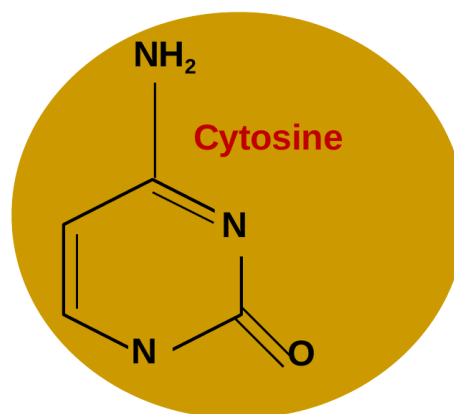
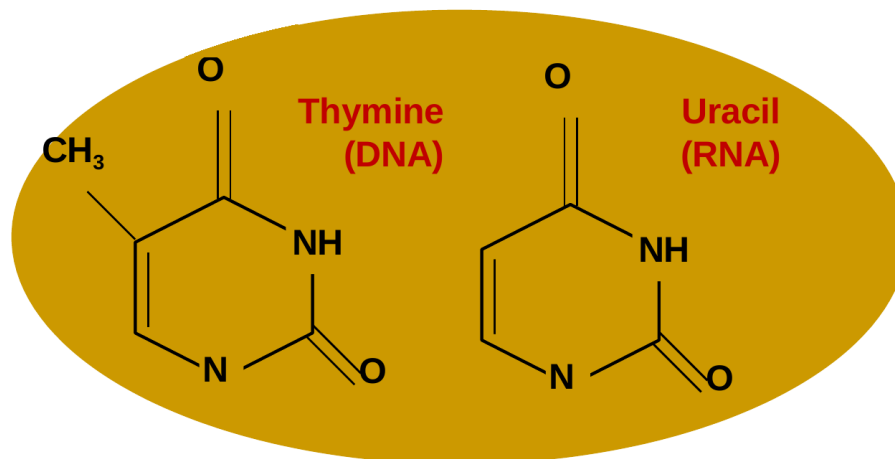


The nitrogenous bases: purines and pyrimidines

Purines



Pyrimidines



Purines · double ring

Adenine (A), Guanine (G)



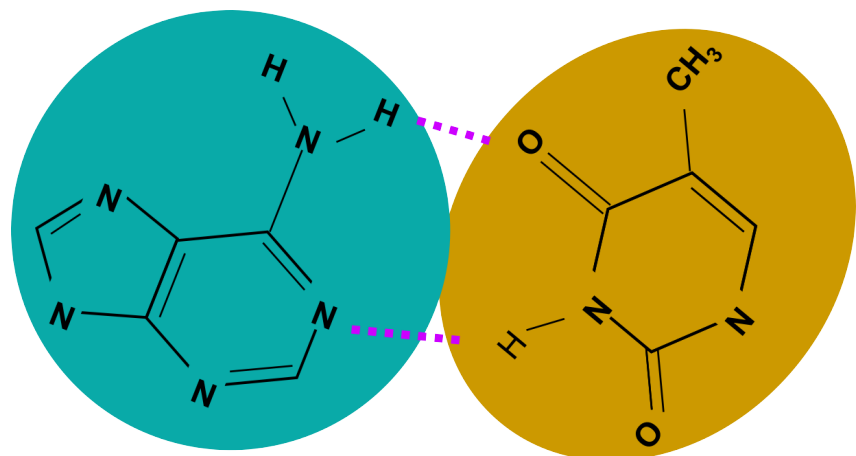
Pyrimidines · single ring

Cytosine (C), Thymine (T)

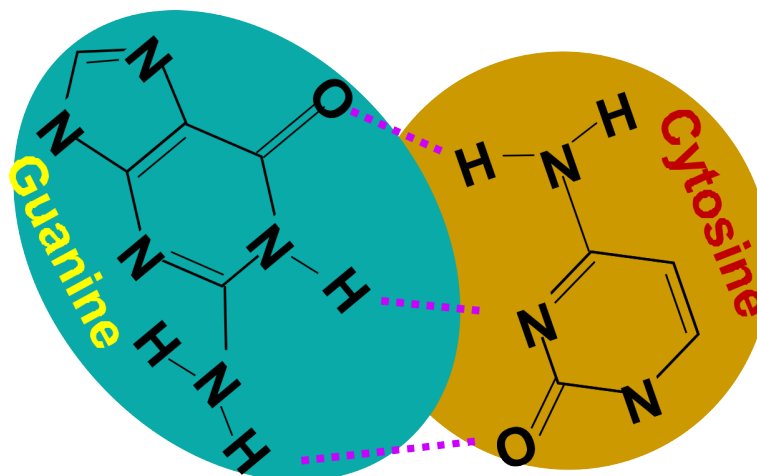
In RNA: uracil (U) instead of T

The hydrogen-bond donors and acceptors on the edge of each base determine which base it can pair with.

Complementarity: $A=T$ and $G\equiv C$

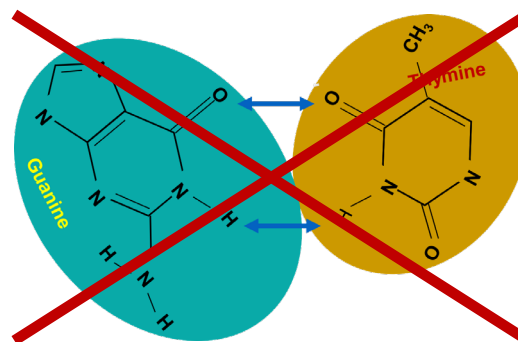
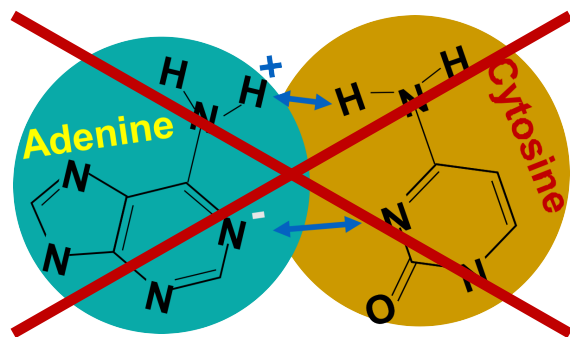


$A = T$ · 2 hydrogen bonds



$G \equiv C$ · 3 hydrogen bonds

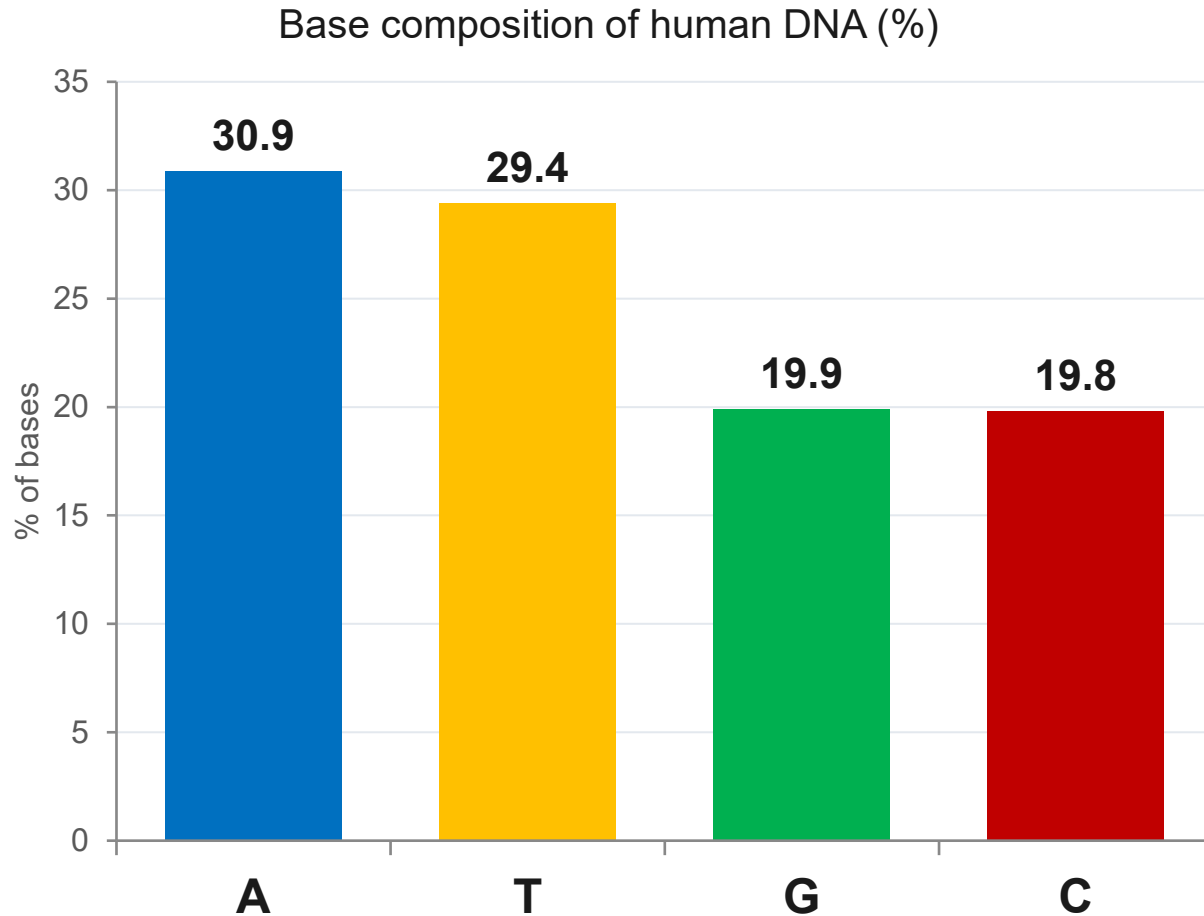
- Purine + pyrimidine → all pairs have the same width
- Weak bonds → the strands separate without breaking covalent bonds
- Many bonds + base stacking → a stable helix
- Mismatches = errors “caught” by DNA repair (Ch. 6)



A-C and G-T mismatches: donor faces donor, acceptor faces acceptor



Chargaff's rules: $A \approx T$ and $G \approx C$



Human DNA data from E. Chargaff's laboratory (early 1950s)

Rule 1

$A \approx T$ and $G \approx C \Rightarrow$ purines = pyrimidines
($A + G = C + T$)

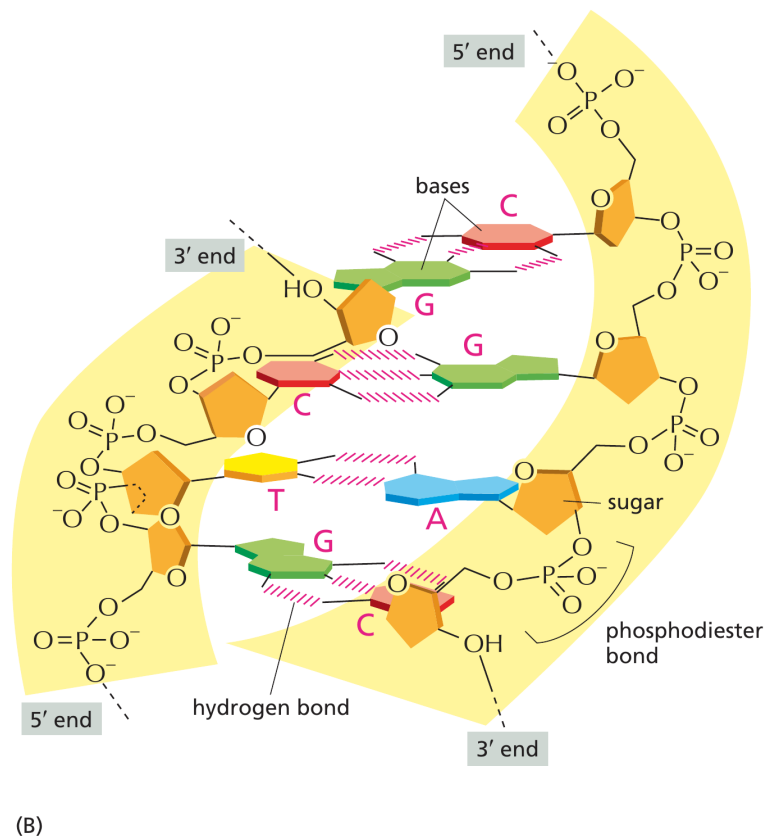
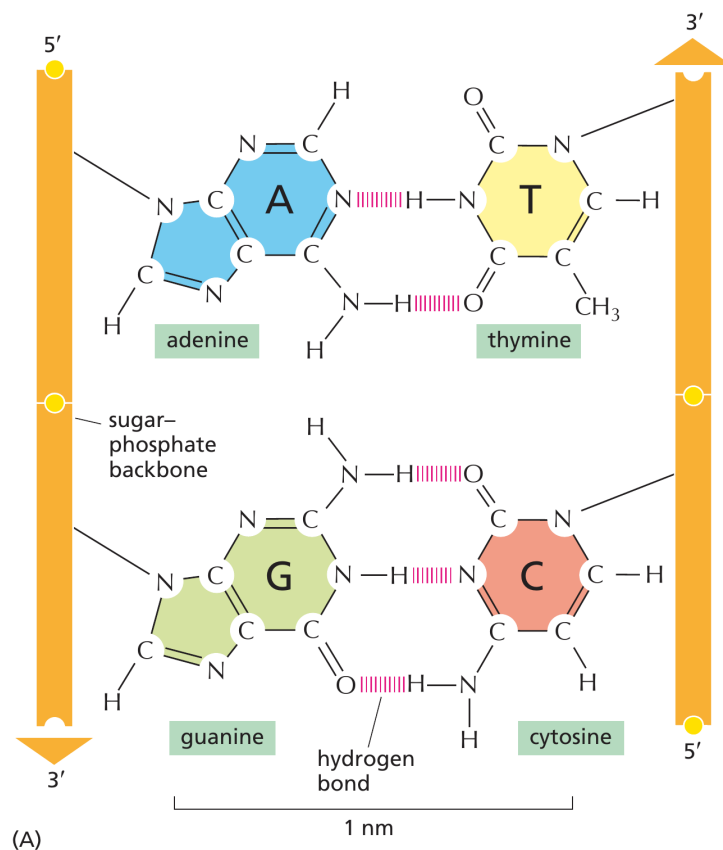
Rule 2

The $(A + T)/(G + C)$ ratio differs between species

The explanation came in 1953

Every A faces a T and every G faces a C, so the amounts must be equal.

Two antiparallel, complementary strands



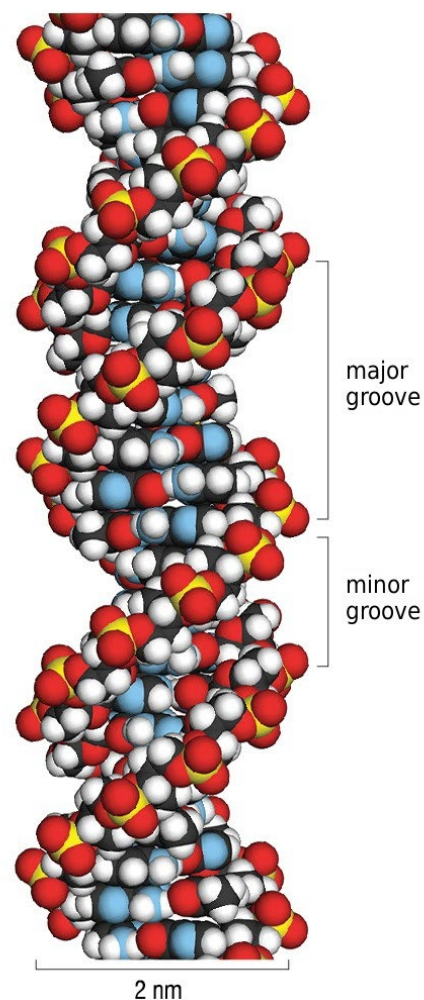
- **Sugar–phosphate backbone** on the outside, in contact with water; bases inside
- **Antiparallel:** one runs 5'→3', the other 3'→5'
- **Complementary:** the sequence of one defines the other

Convention

We always write 5'→3':

5' –ATGC– 3'

The double helix: dimensions and grooves



2 nm

diameter



0.34 nm

per base pair



~10 bp

per turn (3.4 nm)



Right-handed

B-form DNA

In solution: 10.4–10.5 base pairs per turn

Major and minor grooves

In the major groove the edges of the bases are exposed: proteins (e.g. transcription factors) “read” the sequence without opening the helix.

► [Video \(6'\)](#)

[The Structure of DNA · MIT](#)

► [HHMI: Paired DNA Strands](#)

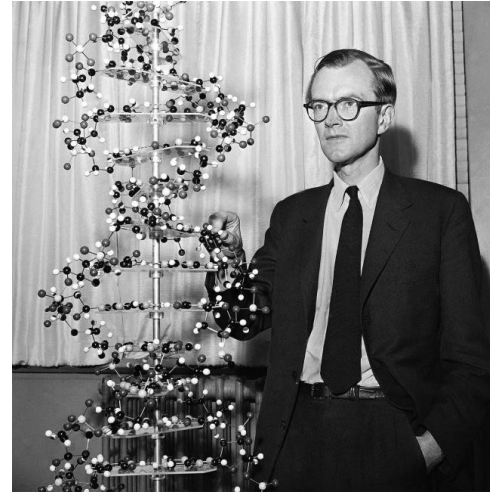
The people behind the double helix



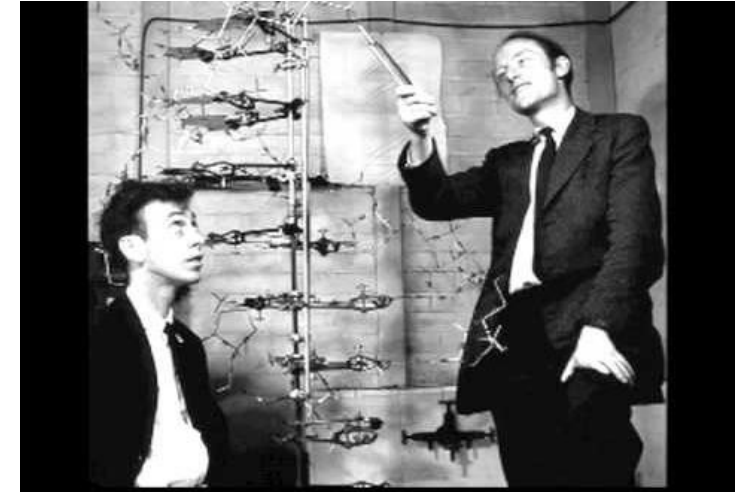
R. Franklin (1920–1958)



"Photo 51", May 1952



Maurice Wilkins



James Watson & Francis Crick

- **1952:** Franklin & Gosling (King's College): X-ray diffraction → a helix and its dimensions
- **25 April 1953:** the Watson & Crick model in *Nature*, in the same issue as papers by Wilkins and by Franklin
- **1962:** Nobel Prize to Watson, Crick and Wilkins; Franklin had died in 1958

Revisiting the story (*Nature*, 2023)

Archival evidence shows that Franklin was an equal member of the “quartet” that solved the structure, not just the source of a photograph.

The structure suggests a mechanism of heredity

devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

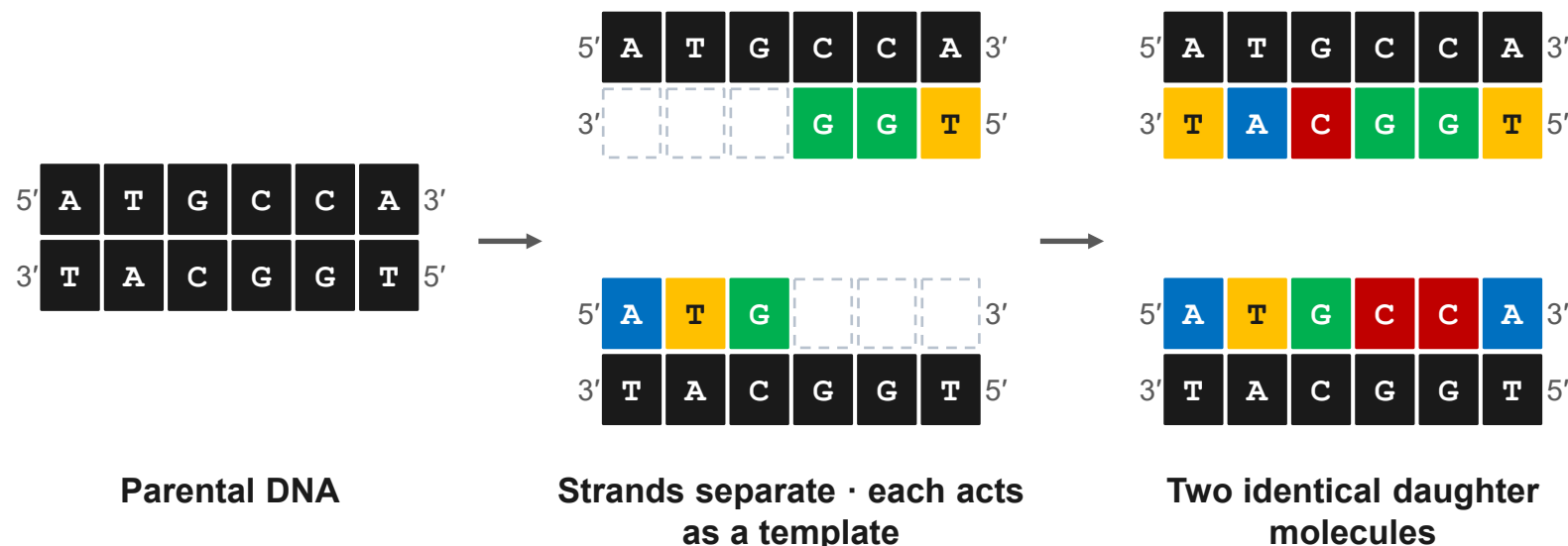
It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

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 inter-atomic 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 King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK

Medical Research Council Unit for the
Study of the Molecular Structure of
Biological Systems,
Cavendish Laboratory, Cambridge.
April 2.



■ old (parental) strand ■ new strand (A↔T, G↔C)

Semiconservative replication (Meselson & Stahl, 1958): each daughter molecule has one old and one new strand. Details: Ch. 6, DNA Replication.

“...immediately suggests a possible copying mechanism for the genetic material.”

Watson & Crick, Nature, 1953

Three quick calculations

a

Write the complementary strand of

5' -ATGCCA-3'

in the 5'→3' direction.

b

A double-stranded DNA contains **30% G**. What percentage is A?

c

Which duplex melts at the higher temperature?

5' -AATTATAT-3'

or

5' -GCCGGCGC-3'

(each paired with its complementary strand)

Complementarity: from the cell to the clinic



IN THE CELL



Replication

Each strand is a template for a new one → heredity



Transcription & translation

DNA → RNA → protein; mRNA codon
↔ tRNA anticodon



Repair

The intact strand guides repair of the damaged one

IN MEDICINE



Diagnosis

PCR, probes, exome/genome sequencing (WES/WGS)



Genome editing

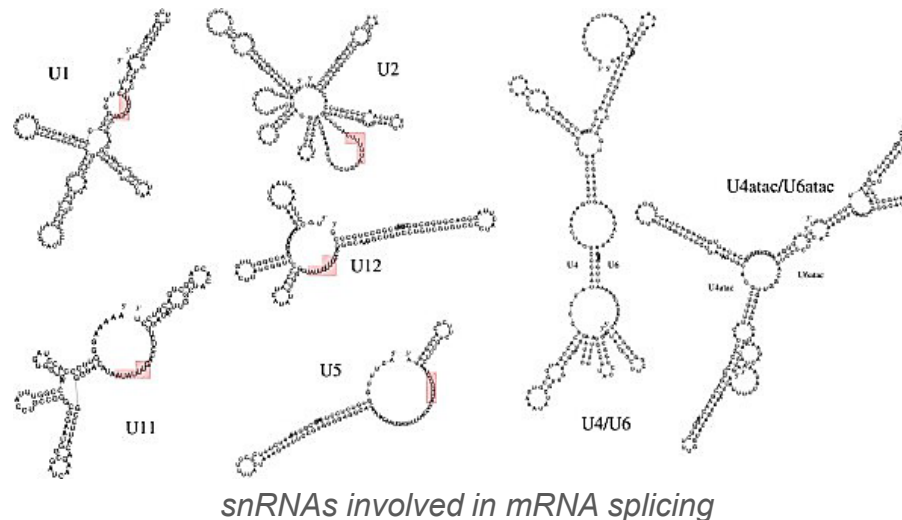
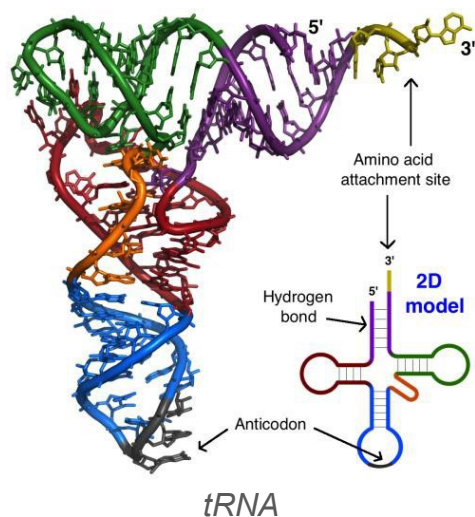
The CRISPR-Cas9 guide RNA finds its target by base pairing



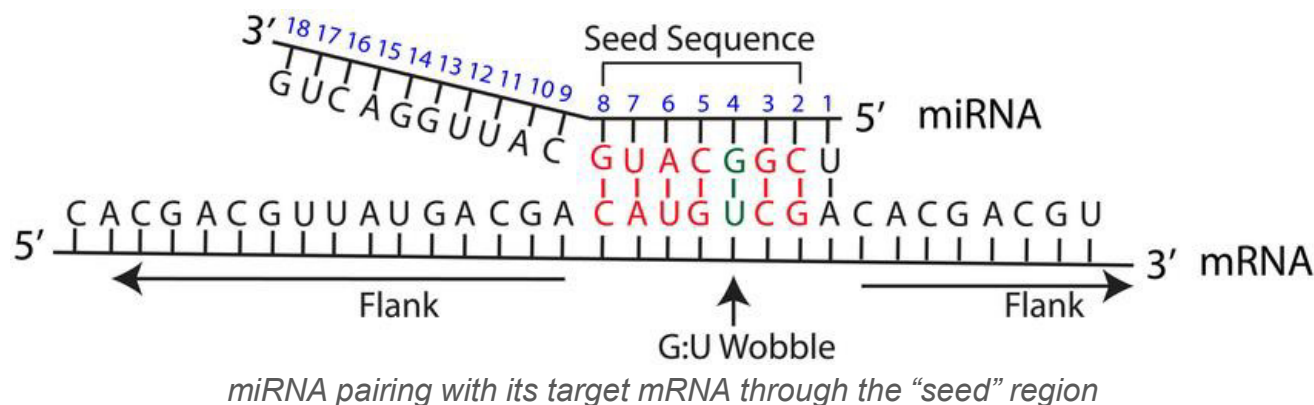
Nucleic acid drugs

Antisense oligonucleotides, siRNAs: bind complementary mRNA

Is double-stranded structure unique to DNA?



V. Ambros & G. Ruvkun · Nobel Prize in Physiology or Medicine 2024 (miRNA)



- RNA is single-stranded but folds on itself: A–U, G–C and G–U pairs
- miRNAs recognize their target mRNA by base pairing (7–8 nucleotides)

The information is written in the sequence

(A) molecular biology is...



(C) • — • — • • — • — • •

(D) 细胞生物学乐趣无穷

(E) TTCGAGCGACCTAACCTATAG

The same phrase in English, as a musical score, in Morse code, in Chinese and in DNA (Alberts et al., Essential Cell Biology, 4th ed., Fig. 5-8)

Morse code: only 2 symbols

• — • — • — • — • — •

= WELCOME

— • — • — • — • — • — •

= MEDICAL

• • — • — • — • — • — •

= STUDENTS

Question #3: How many different “messages” 10 nucleotides long are possible?

4

bases: A, T, G, C

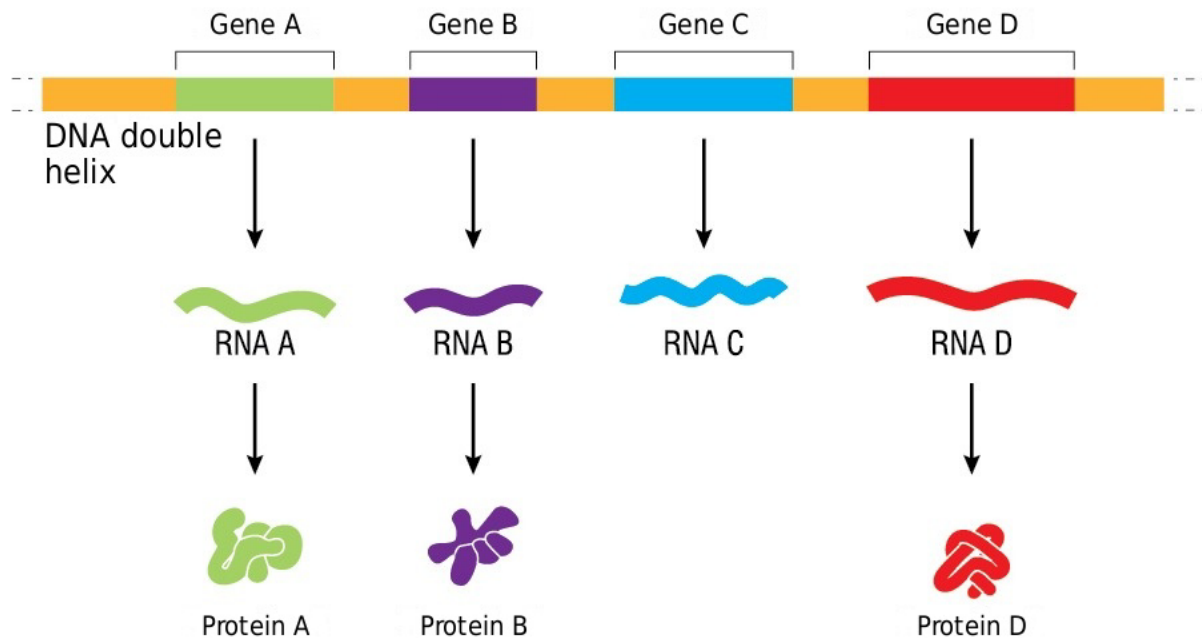
4^n

possible sequences of length n

3.1×10^9

base pairs in the human genome

Gene and genome



Gene: a segment of DNA that carries the information for a functional product (protein or RNA).

Genome: the total genetic information of an organism.

3.055×10^9

base pairs · complete haploid human genome (T2T, 2022)

~20,000

protein-coding genes

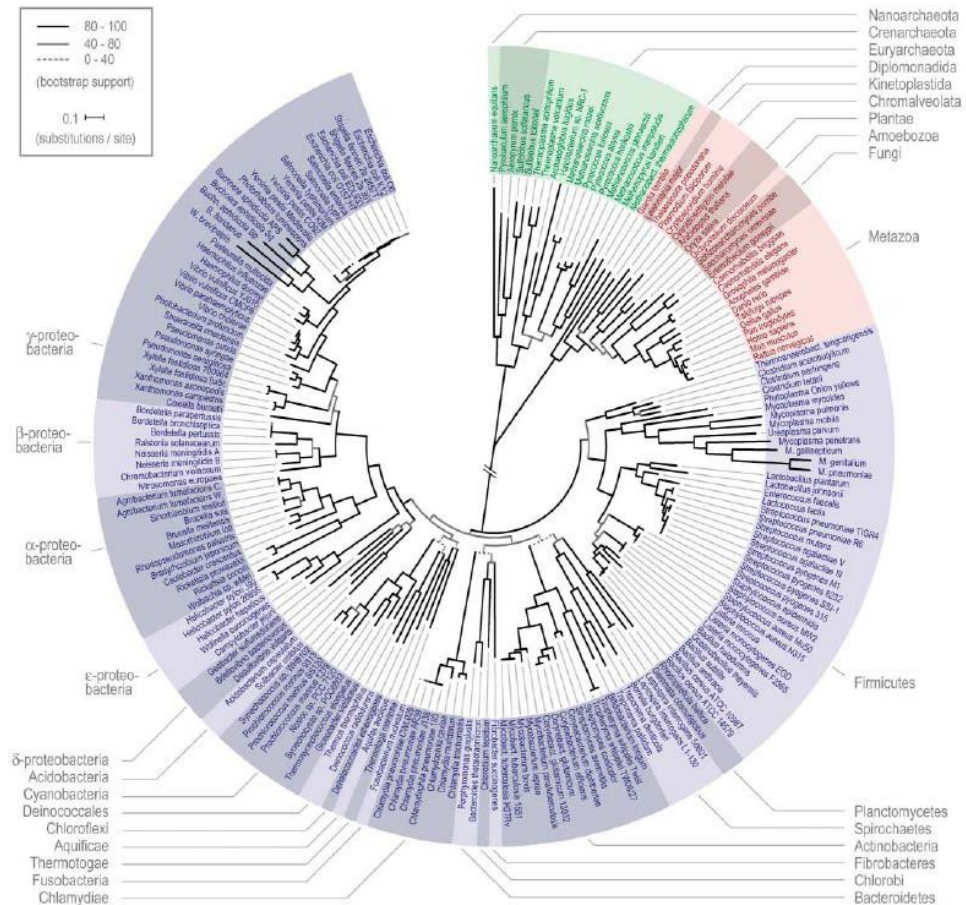
~1.5%

of the genome codes for protein

Back to our patient

HBB (chromosome 11): 3 exons, 2 introns → hemoglobin β chain, 146 amino acids. One base changes → one amino acid changes.

DNA as a record of evolution and health



Svante Pääbo

Nobel Prize 2022: ancient DNA and the genomes of Neanderthals and other extinct hominins

- **Comparing sequences** → **common ancestry**: a tree built from 31 proteins in 191 species
- **Medical relevance**: a ~50 kb segment on chromosome 3, inherited from Neanderthals, is the major genetic risk factor for severe COVID-19 (~16% of Europeans, ~50% in South Asia)
- **Try it yourself**: BLAST in the Bioinformatics lab

Fig. 2. Global phylogeny of fully sequenced organisms. The phylogenetic tree has its basis in a cleaned and concatenated alignment of 31 universal protein families and covers 191 species whose genomes have been fully sequenced (14). Green section, Archaea; red, Eukaryota;

blue, Bacteria. Labels and color shadings indicate various frequently used subdivisions. The branch separating Eukaryota and Archaea from Bacteria in this unrooted tree has been shortened for display purposes.

Summary



- 1 Griffith → Avery → Hershey–Chase: the genetic material is DNA.
- 2 DNA is a polymer of nucleotides with 5'→3' polarity.
- 3 Two antiparallel, complementary strands ($A=T$, $G\equiv C$) form a right-handed double helix.
- 4 Complementarity enables replication, expression and repair, and underpins molecular diagnosis and therapy.
- 5 The information lies in the linear sequence: gene → RNA → protein.



I. True or false? Justify briefly (1–3 sentences)

- A** In the Hershey & Chase experiment, the ^{35}S radioactivity was found mainly in the bacterial pellet after centrifugation.
- B** If a double-stranded DNA molecule contains 20% adenine, it contains 30% guanine.
- C** A G–C-rich DNA segment melts at a higher temperature than an A–T-rich segment of the same length.
- D** The two strands of the double helix are parallel: both run 5'→3' in the same direction.
- E** The genetic information of DNA is encoded in its sugar–phosphate backbone.

II. Match the terms in pairs and justify

Frederick Griffith

^{32}P

Deoxyribonuclease (DNase)

Rosalind Franklin

Purines = pyrimidines

Transformation

Hershey & Chase

Avery, MacLeod & McCarty

X-ray diffraction

Erwin Chargaff

Each term is used once · form 5 pairs · justify each pair in 1–3 sentences

III. Clinical scenario: sickle cell disease

In a patient with sickle cell disease, sequencing of the HBB gene revealed the change below (coding strand; the altered nucleotide is underlined):

5' -ATG GTG CAC CTG ACT CCT GAG GAG AAG-3' normal HBB

5' -ATG GTG CAC CTG ACT CCT GTG GAG AAG-3' mutant HBB

- Write the complementary strand of the normal segment in the 5'→3' direction.
- How does the number of hydrogen bonds change at the mutated position? Explain.
- Explain why a change in a single base pair can alter the function of a protein.
- Exa-cel therapy uses CRISPR-Cas9 with a guide RNA. Which property of nucleic acids does target recognition rely on?

For study and review



Animations & videos

[DNA from the Beginning 17: Griffith & Avery](#)

[DNA from the Beginning 18: Hershey & Chase](#)

[DNA from the Beginning 19: the double helix](#)

[HHMI BioInteractive: Paired DNA Strands](#)

[MIT: The Structure of DNA \(6'\)](#)

Reading & references

Alberts et al. Essential Cell Biology, 6th ed. (2023), Ch. 5 · Cooper & Adams, The Cell: A Molecular Approach, 9th ed. · Alberts et al. Molecular Biology of the Cell, 7th ed., Ch. 4

Griffith 1928 J Hyg · Avery et al. 1944 J Exp Med · Hershey & Chase 1952 J Gen Physiol · Watson & Crick 1953 Nature

Cobb & Comfort 2023 Nature · Nurk et al. 2022 Science (T2T) · Liao et al. 2023 Nature (pangenome)

Frangoul et al. 2024 NEJM (exa-cel) · Musunuru et al. 2025 NEJM · Zeberg & Pääbo 2020 Nature · Ciccarelli et al. 2006 Science



Next lecture

Structure and organization of eukaryotic chromosomes: how ~2 m of DNA fits into a nucleus a few micrometers across