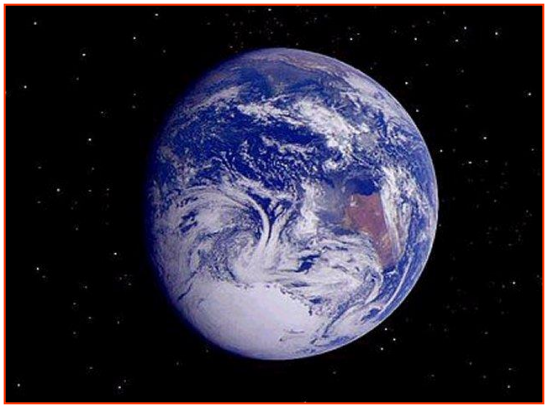
Basic Features of Eukaryotic Genes

Dr.G.MATHAN
Assistant Professor
Department of Biomedical Science
Bharathidasan University
Tiruchirappalli, Tamil Nadu



4.54 billion years old Earth

Living cells evolving and diversifying for over 3.5 billion years

Three Major Division of Living World

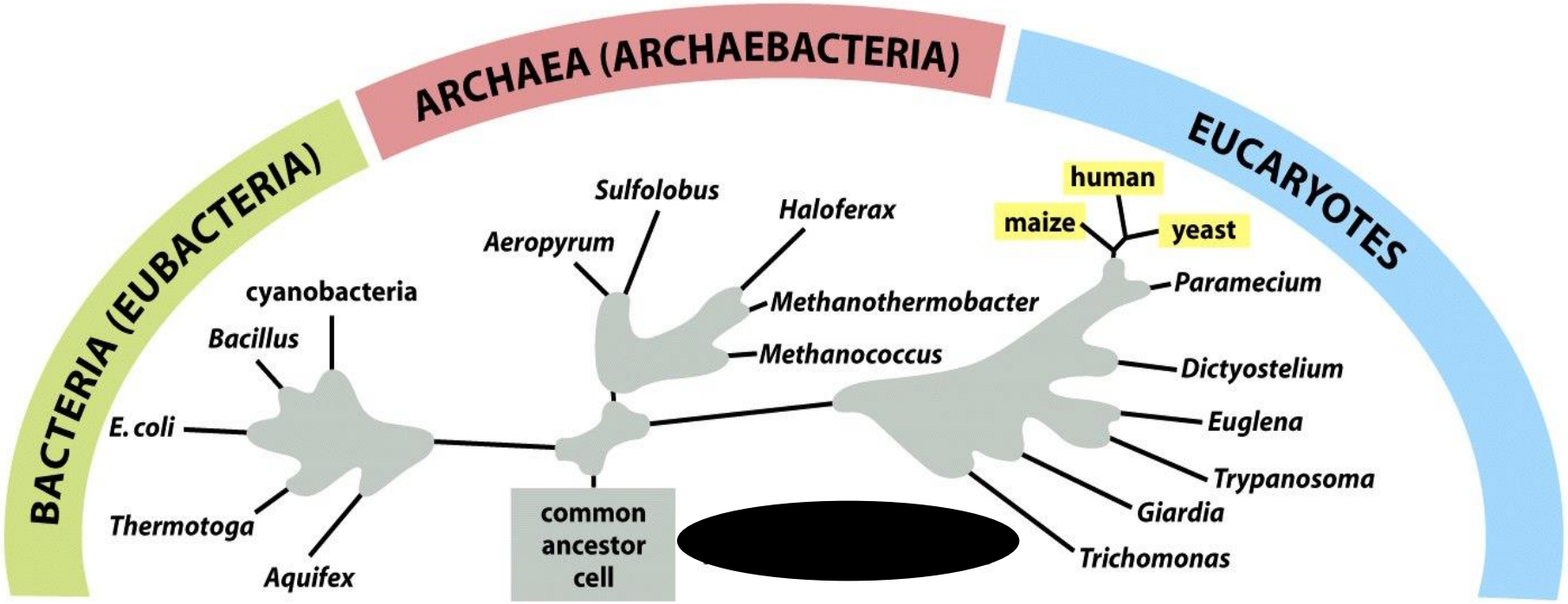


Figure 1-21 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Figure 1-22 Molecular Biology of the Cell 5/e (© Garland Science 2008)

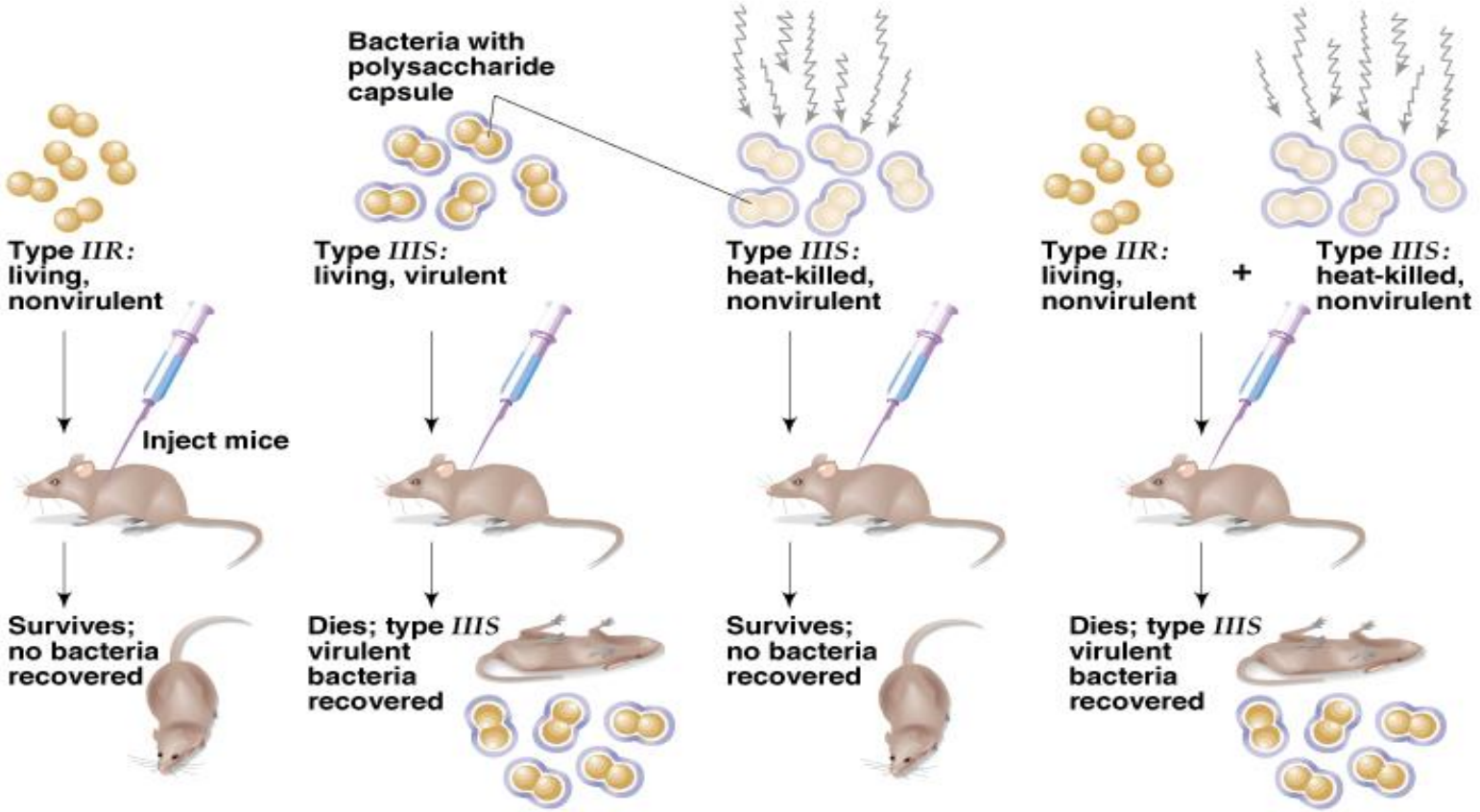
Search for the genetic material

Timeline of events:

- **1890** — **Weismann - substance in the cell nuclei controls development.**
- **1900** — **Chromosomes shown to contain hereditary information, later shown to be composed of protein & nucleic acids.**
- **1928** — **Griffith's Transformation Experiment**
- **1944** — **Avery's Transformation Experiment**
- **1953** — **Hershey-Chase Bacteriophage Experiment**
- **1953** — **Watson & Crick propose double-helix model of DNA**
- **1956** — **First demonstration that RNA is viral genetic material.**

Frederick Griffith's Transformation Experiment - 1928

"transforming principle" demonstrated with *Streptococcus pneumoniae*

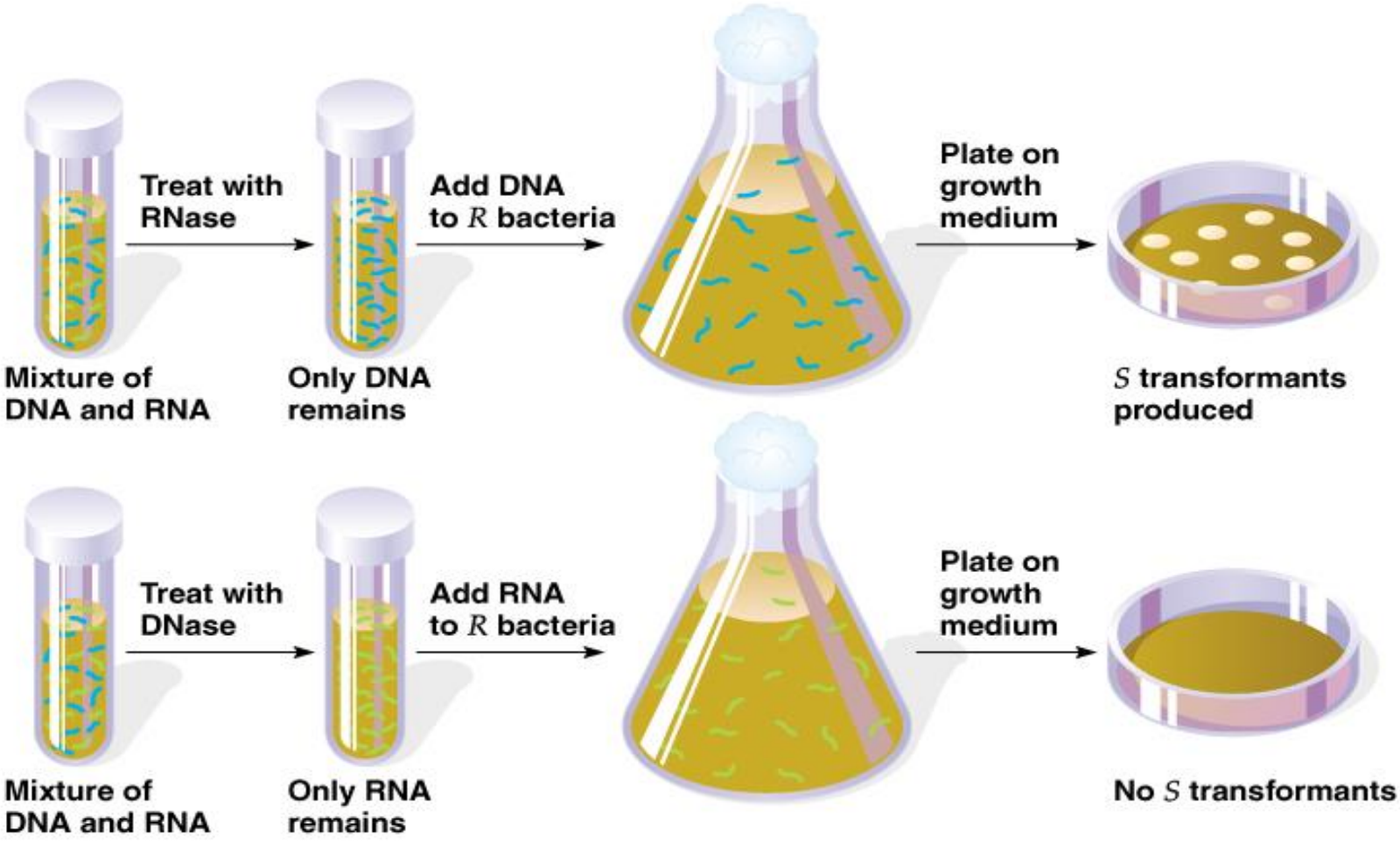


Griffith hypothesized that the transforming agent was a "IIIS" protein. But this was only a guess, and Griffith turned out to be wrong.

The III-S strain DNA contains the genes that form the protective polysaccharide capsule. Equipped with this gene

Oswald T. Avery's Transformation Experiment - 1944

Determined that "IIIS" DNA was the genetic material responsible for Griffith's results (not RNA).



Hershey-Chase Bacteriophage Experiment - 1953

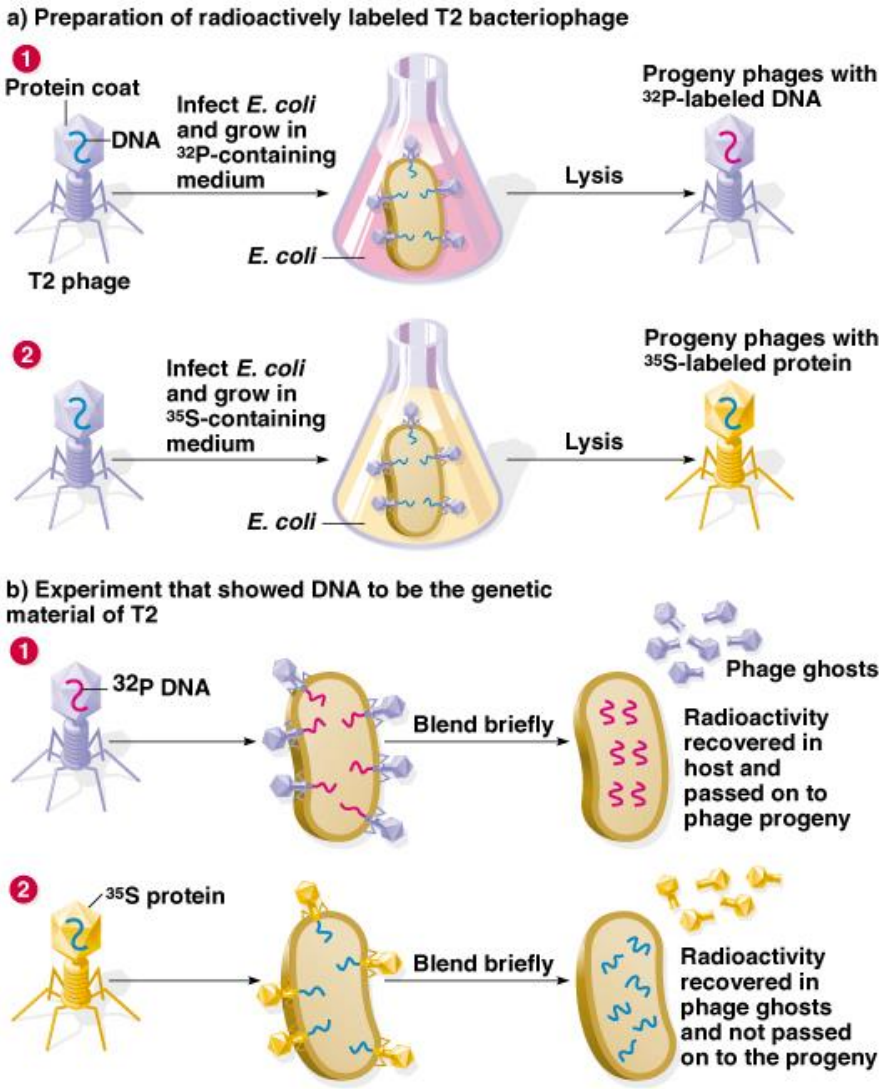
1. T2 bacteriophage is composed of DNA and proteins:

2. Set-up two replicates:

   - Label DNA with ^{32}P
   - Label Protein with ^{35}S

3. Infected *E. coli* bacteria with two types of labeled T2

4. ^{32}P is discovered within the bacteria and progeny phages, whereas ^{35}S is not found within the bacteria but released with phage hosts.



Alfred Hershey

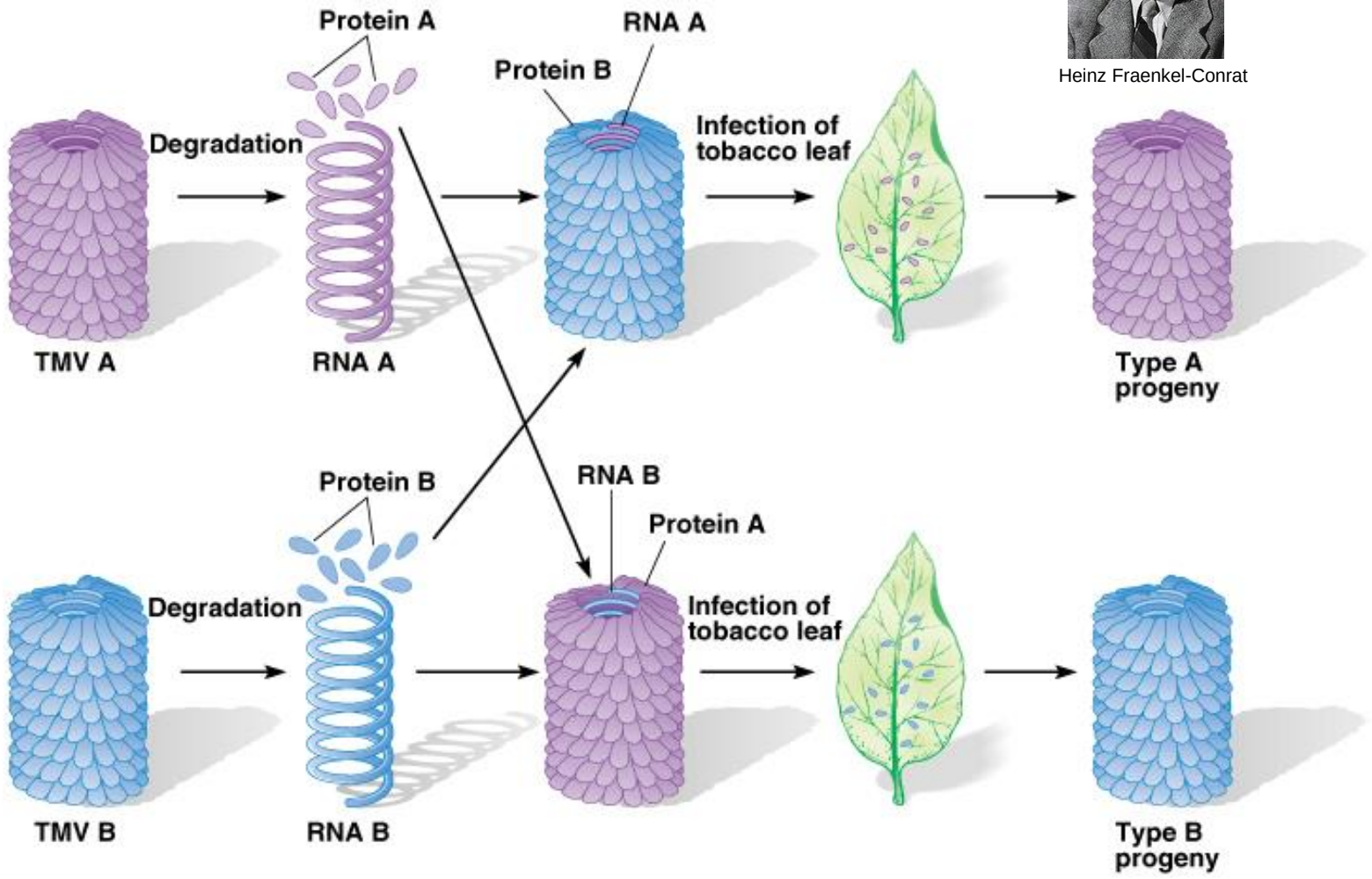
Courtesy of Cold Spring Harbor Laboratory Archives. Noncommercial, educational use only.

Sulfur is present in the amino acids cysteine and methionine, but not in DNA

Gierer & Schramm Tobacco Mosaic Virus (TMV) Experiment - 1956
Fraenkel-Conrat & Singer – 1957, Max Lauffer, David Trkule and Anne Buzzell

Demonstrated that RNA is the genetic material of TMV.


Heinz Fraenkel-Conrat



Conclusions about these early experiments:

Griffith 1928 & Avery 1944:

DNA (not RNA) is transforming agent.

Hershey-Chase 1953:

DNA (not protein) is the genetic material.

Gierer & Schramm 1956/Fraenkel-Conrat & Singer 1957:

RNA (not protein) is genetic material of some viruses, but no known prokaryotes or eukaryotes use RNA as their genetic material.

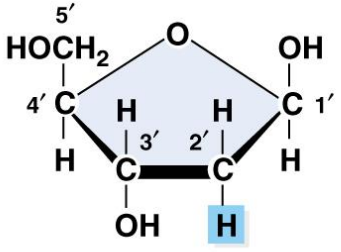
Alfred Hershey
Nobel Prize in Physiology or Medicine
1969

Nucleic Acids

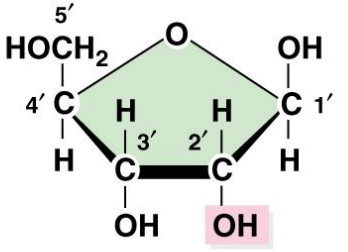
Nucleotide = monomers that make up DNA and RNA

<u>Three components</u>

1. Pentose (5-carbon) sugar
DNA = deoxyribose
RNA = ribose
(compare 2′ carbons)



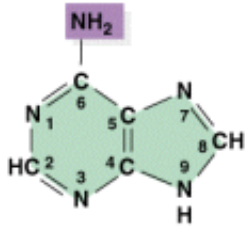
Deoxyribose



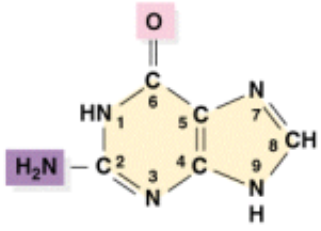
Ribose

2. Nitrogenous base

<u>Purines</u>
Adenine
Guanine

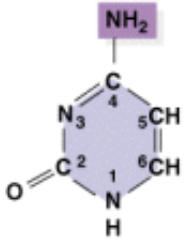


Adenine (A)

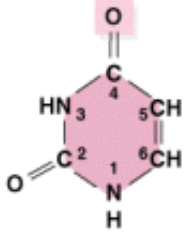


Guanine (G)

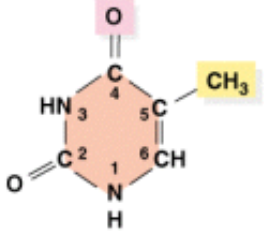
<u>Pyrimidines</u>
Cytosine
Thymine (DNA)
Uracil (RNA)



Cytosine (C)

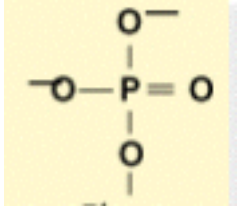


Uracil (U)
(found in RNA)



Thymine (T)
(found in DNA)

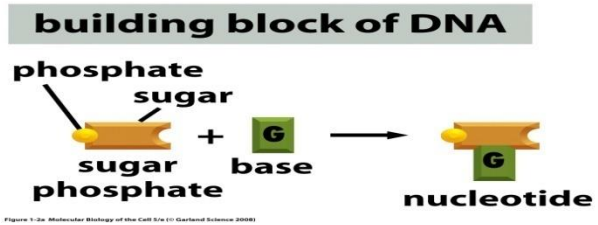
3. Phosphate group attached to 5′ carbon

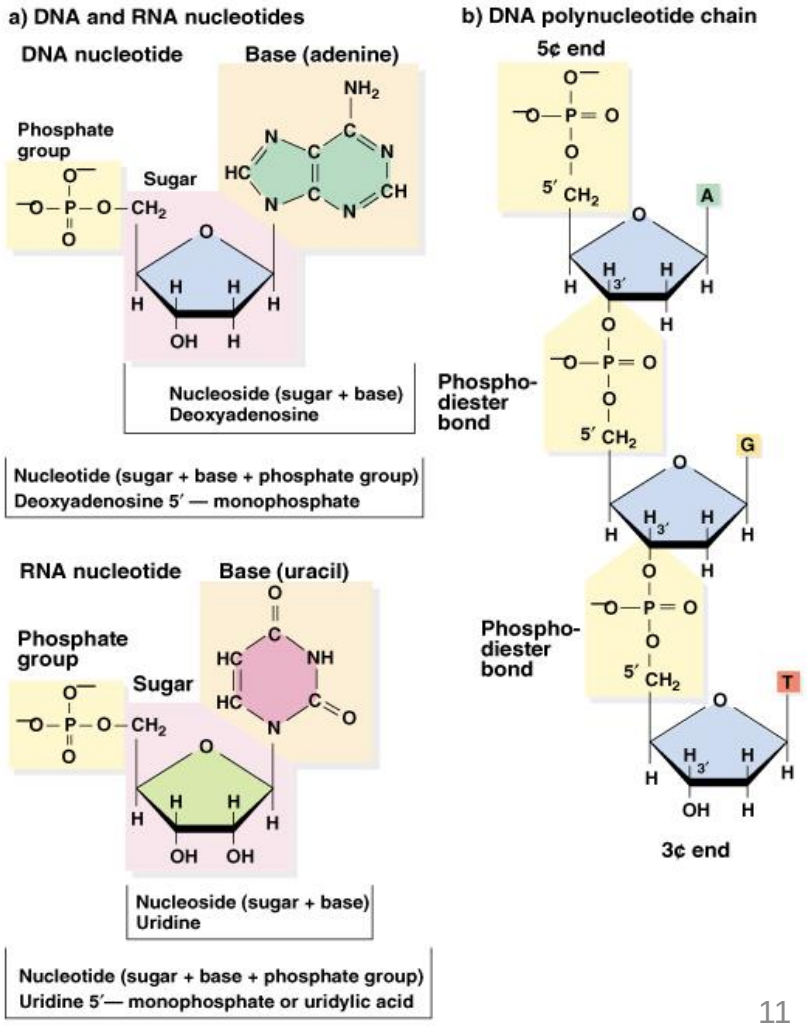


Phosphodiester bond

Covalent bond between the phosphate group (attached to 5’ carbon) of one nucleotide and the 3’ carbon of the sugar of another nucleotide.

Nucleotides are linked by phosphodiester bonds to form polynucleotides.





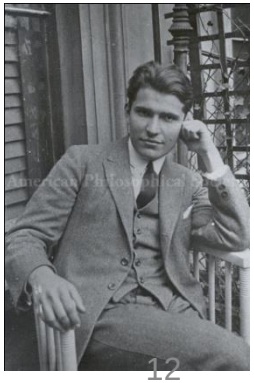
Structure of DNA

James D. Watson/Francis H. Crick 1953 proposed the Double Helix Model based on two sources of information:

1. **Base composition studies of Erwin Chargaff**

   - **indicated double-stranded DNA consists of ~50% purines (A,G) and ~50% pyrimidines (T, C)**
   - **amount of A = amount of T and amount of G = amount of C (Chargraff's rules)**
   - **%GC content varies from organism to organism**

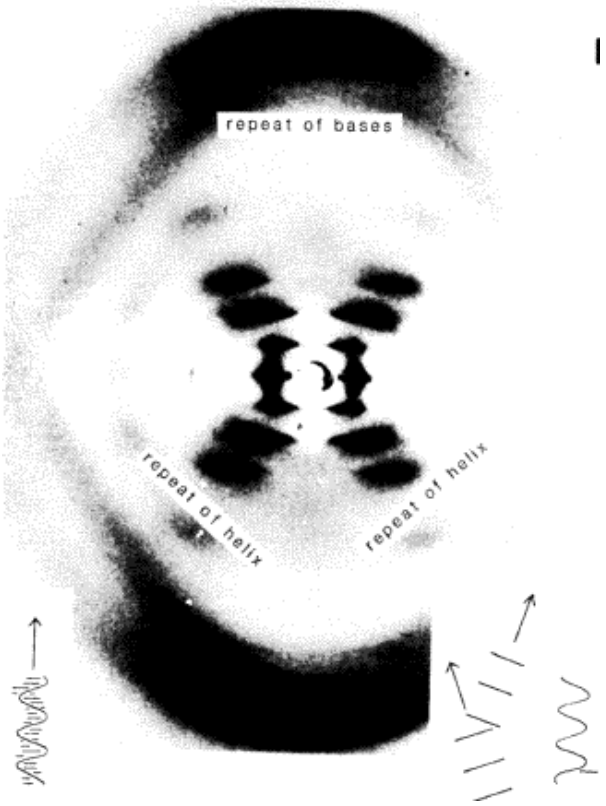
Examples:	**%A**	**%T**	**%G**	**%C**	**%GC**
Homo sapiens	31.0	31.5	19.1	18.4	37.5
Zea mays	25.6	25.3	24.5	24.6	49.1
Drosophila	27.3	27.6	22.5	22.5	45.0
Aythya americana	25.8	25.8	24.2	24.2	48.4



Structure of DNA

James D. Watson/Francis H. Crick 1953 proposed the Double Helix Model based on two sources of information:

2. X-ray diffraction studies by Rosalind Franklin & Maurice Wilkins



Conclusion-DNA is a helical structure with distinctive regularities, 0.34 nm & 3.4 nm.

She died in 1958 at the age of 37 of ovarian cancer.

Double Helix Model of DNA: Six main features

1. **Two polynucleotide chains wound in a right-handed (clockwise) double-helix.**

2. **Nucleotide chains are anti-parallel: 5′ → 3′**
   3′ ← 5′

3. **Sugar-phosphate backbones are on the outside of the double helix, and the bases are oriented towards the central axis.**

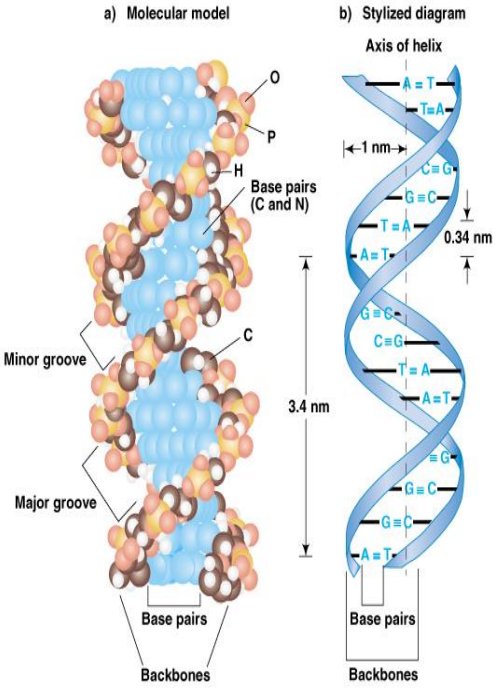
4. **Complementary base pairs from opposite strands are bound together by weak <u>hydrogen bonds</u>.**

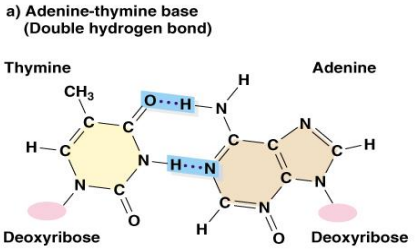
   A pairs with T (<u>2 H-bonds</u>), and G pairs with C (<u>3 H-bonds</u>).

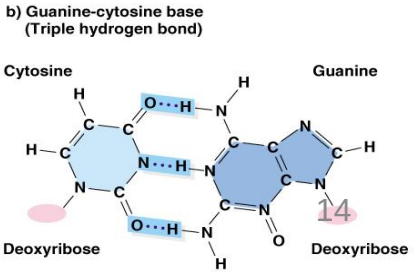
   5′-TATTCCGA-3′
   3′-ATAAGGCT-5′

5. **Base pairs are 0.34 nm apart. One complete turn of the helix requires 3.4 nm (10 bases/turn).**

6. **Sugar-phosphate backbones are not equally-spaced, resulting in major and minor grooves.**







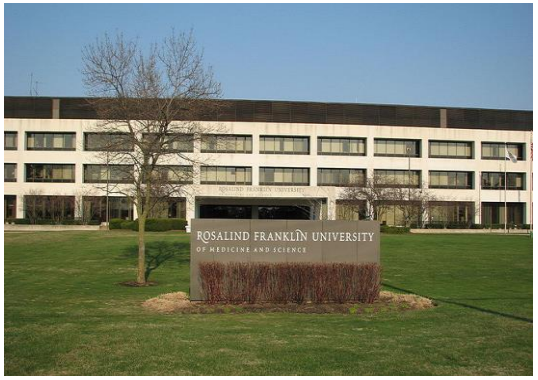
1962: Nobel Prize in Physiology and Medicine

1962: Nobel Prize in Physiology and Medicine

James D. Watson

Francis H. Crick

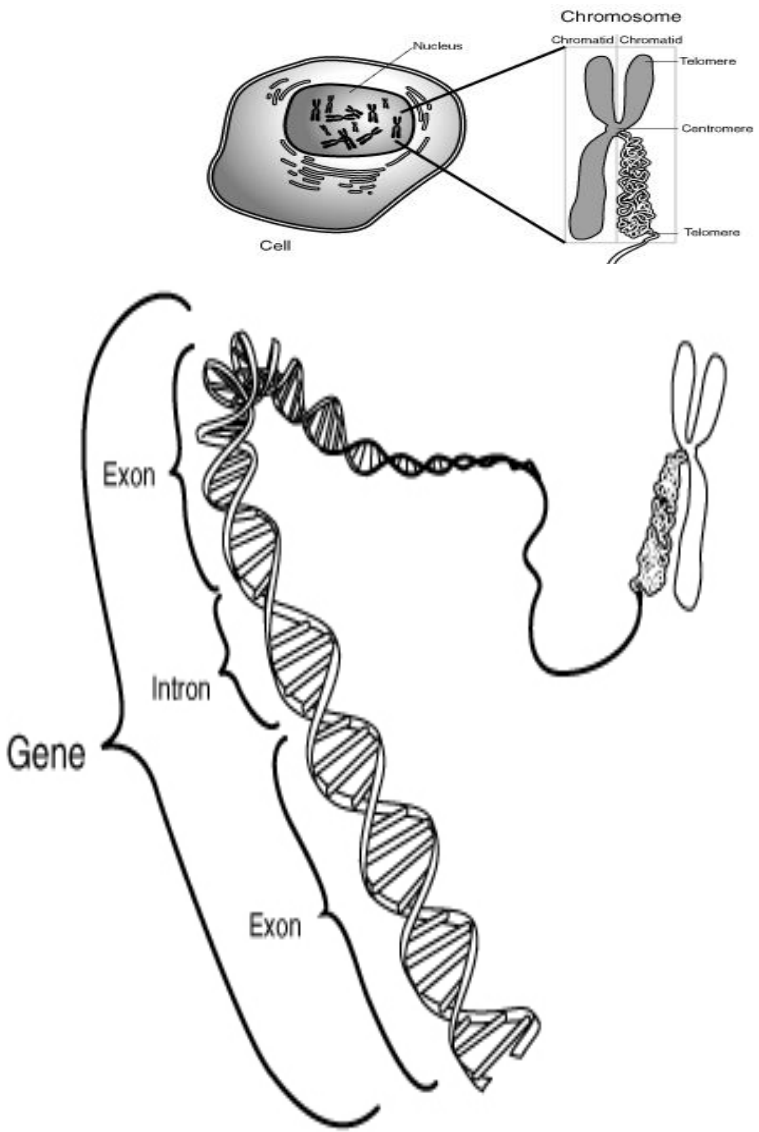
Maurice H. F. Wilkins



North Chicago, Illinois,USA

What about?
Rosalind Franklin

Definitions of the gene



- **The herditary nature of living organism is defined by "Genome".- Consist of a long sequence of nucleic acid –Provides entire genetic material of an individual.**

- **A modern working definition of a gene is "a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and or other functional sequence regions ".**

The Genetic Code

Codon: set of triplet nucleotides that specifies a particular amino acid

Reading frame: the series of nucleotides read in sets of 3 (codon).Only one reading frame is translated.

ORF: Reading frame consisits of triplet representing amino acids.

Start codon: the codon (AUG) used to signify the start of translation

Stop codons: 3 codons (UUA, UGA, UAG) in the genetic code used to terminate translation

TABLE 15.1 The Genetic Code

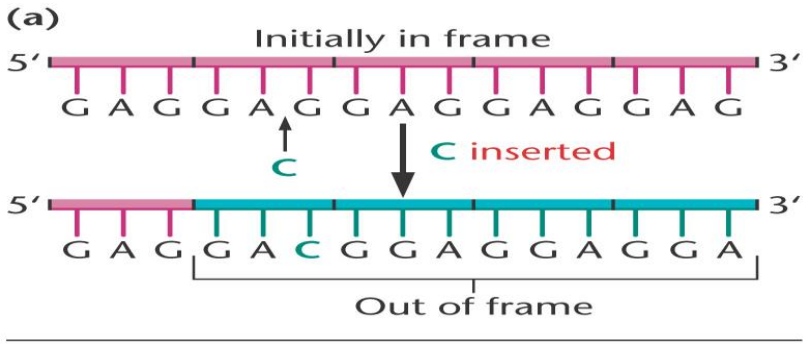
First Letter	SECOND LETTER U		C		A		G		Third Letter
U	UUU	Phe Phenylalanine	UCU	Ser Serine	UAU	Tyr Tyrosine	UGU	Cys Cysteine	U
	UUC		UCC		UAC		UGC		G
	UUA	Leu Leucine	UCA		UAA	"Stop"	UGA	"Stop"	A
	UUG		UCG		UAG	"Stop"	UGG	Trp Tryptophan	G
C	CUU	Leu Leucine	CCU	Pro Proline	CAU	His Histidine	CGU	Arg Arginine	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Gln Glutamine	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Ile Isoleucine	ACU	Thr Threonine	AAU	Asn Asparagine	AGU	Ser Serine	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lys Lysine	AGA	Arg Arginine	A
	AUG	Met Methionine; "Start"	ACG		AAG		AGG		G
G	GUU	Val Valine	GCU	Ala Alanine	GAU	Asp Aspartate	GGU	Gly Glycine	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glu Glutamate	GGA		A
	GUG		GCG		GAG		GGG		G

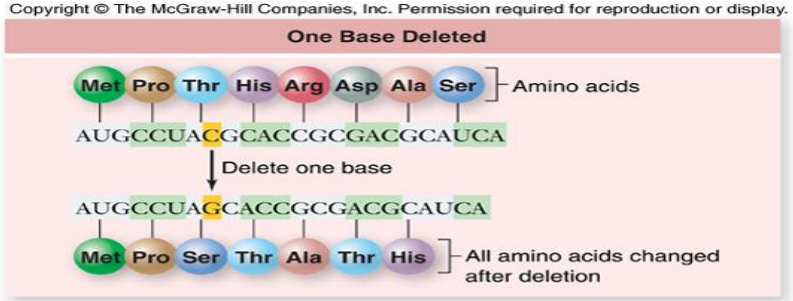
A codon consists of three nucleotides read in the sequence shown. For example, ACU codes for threonine. The first letter, A, is in the First Letter column; the second letter, C, is in the Second Letter column; and the third letter, U, is in the Third Letter column. Each of the mRNA codons is recognized by a corresponding anticodon sequence on a tRNA molecule. Many amino acids are specified by more than one codon. For example, threonine is specified by four codons, which differ only in the third nucleotide (ACU, ACC, ACA, and ACG).

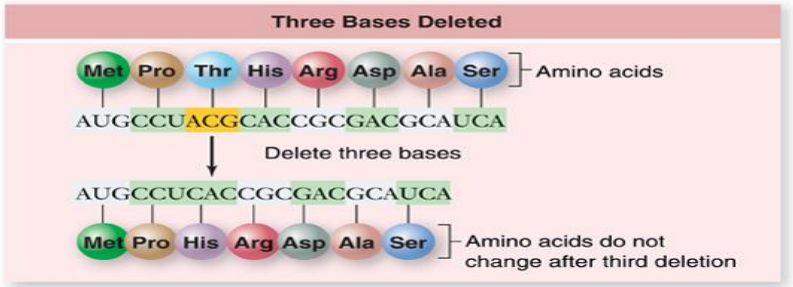
The amino acids encoded by all 64 possible codons were determined.

Triplet Code

- Frameshifts: Inserts or delets by single base will change triplet sets for entire subsequent sequence
- Single nucleotide insertions compensate for single nucleotide deletions
- Two single nucleotide insertions still give frameshift
- A total of 3 added or deleted nucleotides leave code in frame







The flow of Genetic Information

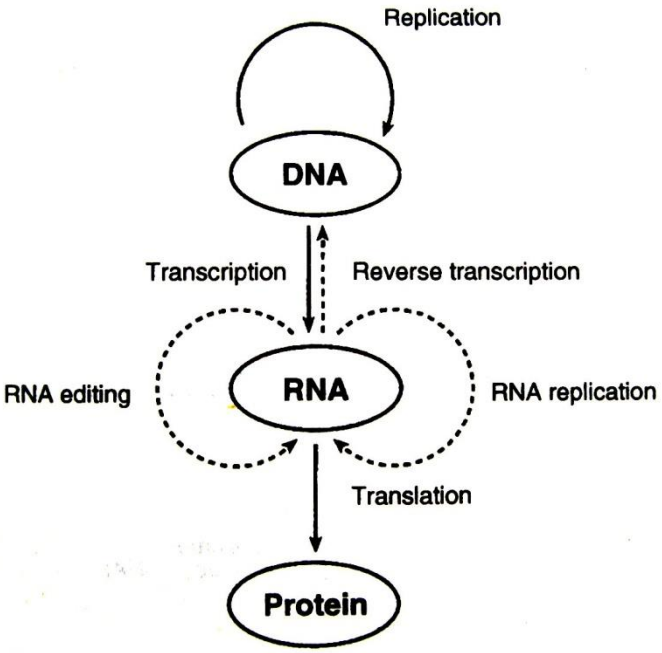
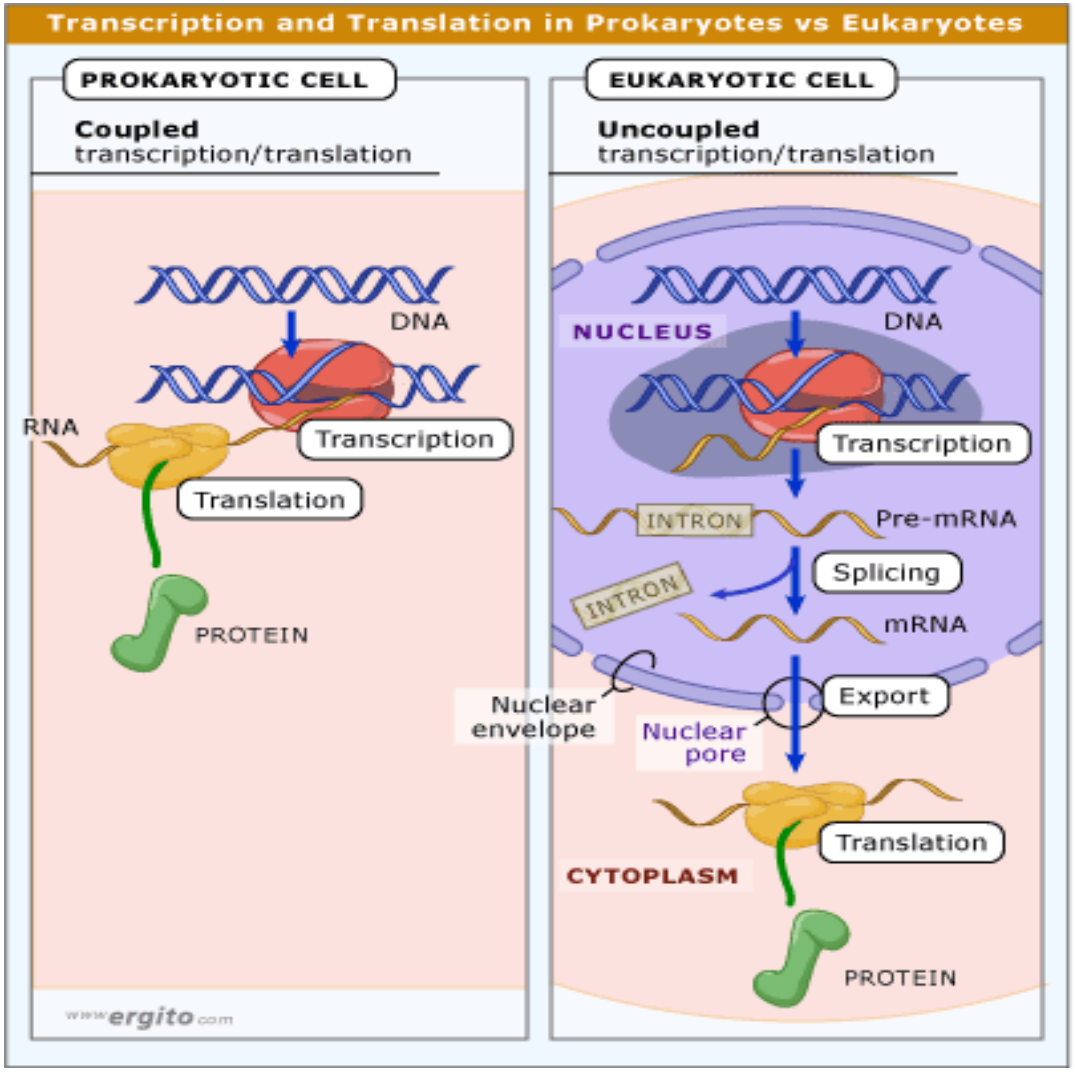
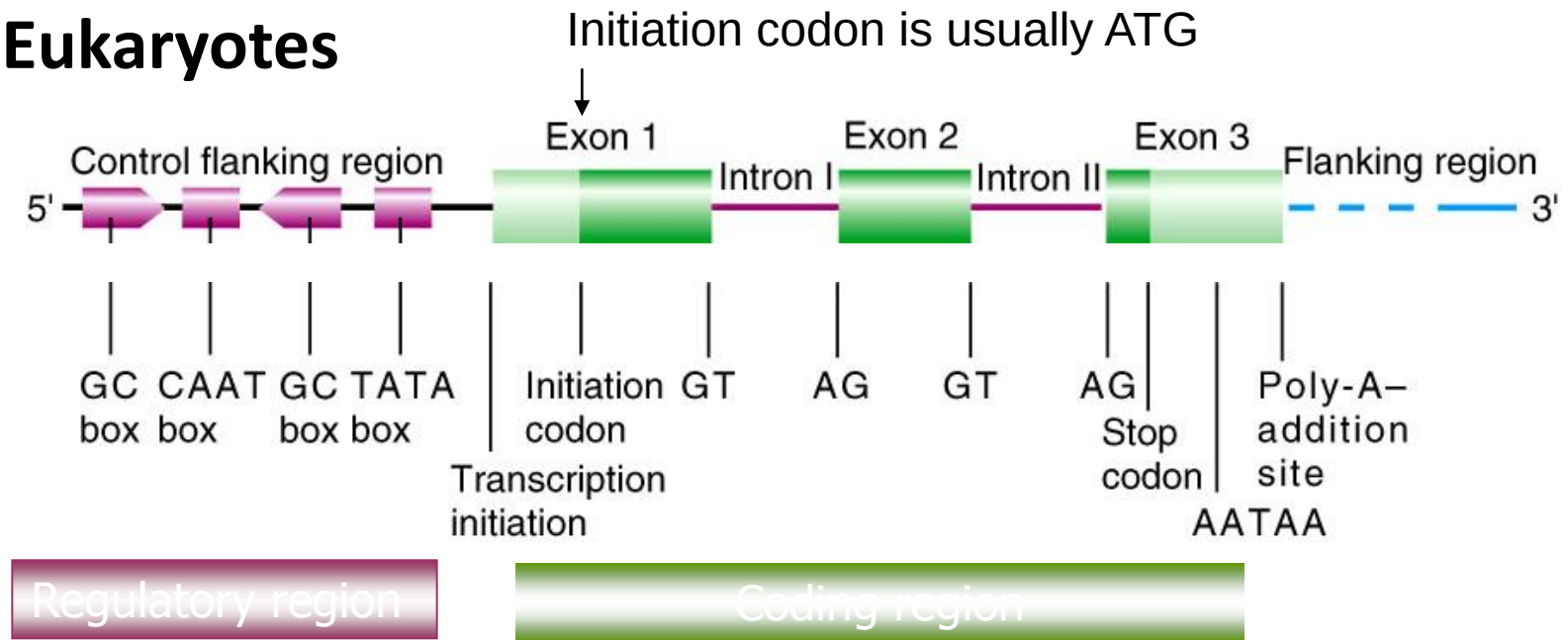


Fig. 1. The flow of genetic information.



Basic gene structure

Eukaryotes



What is a transcription factor?

A transcription factor is a protein that regulates transcription after nuclear translocation by specific interaction with DNA or by stoichiometric interaction with a protein that can be assembled into a sequence-specific DNA-protein complex.

Three classes of transcription factors

1. **General transcription factors (GTFs):**
   Synthesis of all mRNAs
   Select the transcriptional initiation site and delivering Pol II

2. **Upstream transcription factors:**
   Bind to specific DNA sequences upstream of the initiation site so as to stimulate or repress transcriptional initiation by GTF-complexed Pol II. The binding of upstream factors to DNA is unregulated.

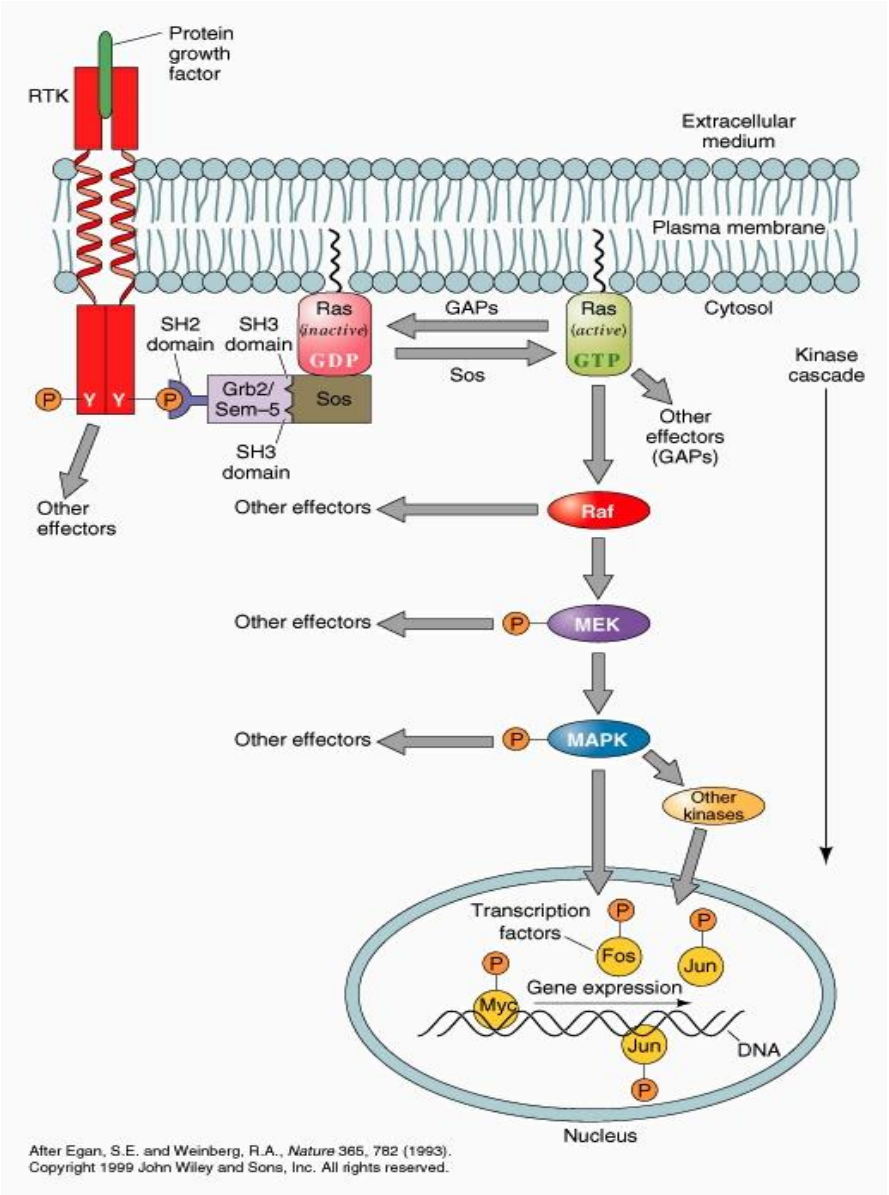
3. **Inducible transcriptional factors.**
   Proteins function similar to upstream transcriptional factors but must be activated (or inhibited), either by phosphorylation or by specific ligands, in order to bind to their target DNA site and influence transcriptional initiation.

TF-CELL SIGNALLING TARGETS

Whenever cells need to response to an extracellular signal such as a hormone, the response is mediated by a change in gene expression that comes about, most often as the result of a change in the phosphorylation state of a transcription factor.

For example,

Growth factor binding to its receptor catalyses the autophosphorylation of a tyrosine residue in the cytoplasmic domain of the receptor. This, in turn is recognised by the SH2 (***Src*** homology 2) domain of a cytoplasmic response protein which through its further interactions activates the Ras protein. Ras is a G protein that is active when GTP is bound but inactive when GDP is bound. Ras then activates a series of kinases until, finally, one of these migrates to the nucleus where it phosphorylates a transcription factor such as Fos, Jun or Myc.



G-guanosine nucleotide-binding proteins

Thanks for your attention!

Acknowledgement

- The Presentation is being used for educational and non commercial purpose
- Thanks are due to all those original contributors and entities whose pictures used for making this presentation.