Introduction to....

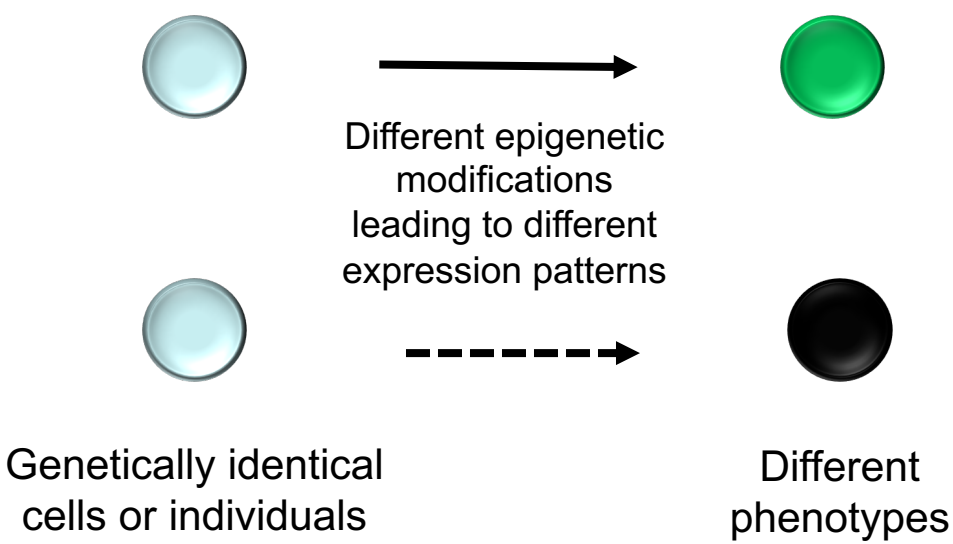
**Your behaviors and environment can induce changes influencing the way your genes work**

# What does “epigenetics” mean?

**Literally**, epigenetics means above, or on top of, genetics.

**Practically**, epigenetics describes phenomena in which genetically identical cells or organisms express their genes differently, causing phenotypic differences.

Riggs et al. 1996



# Gene expression

Genetic changes can influence which gene is expressed and in turns which protein is made.

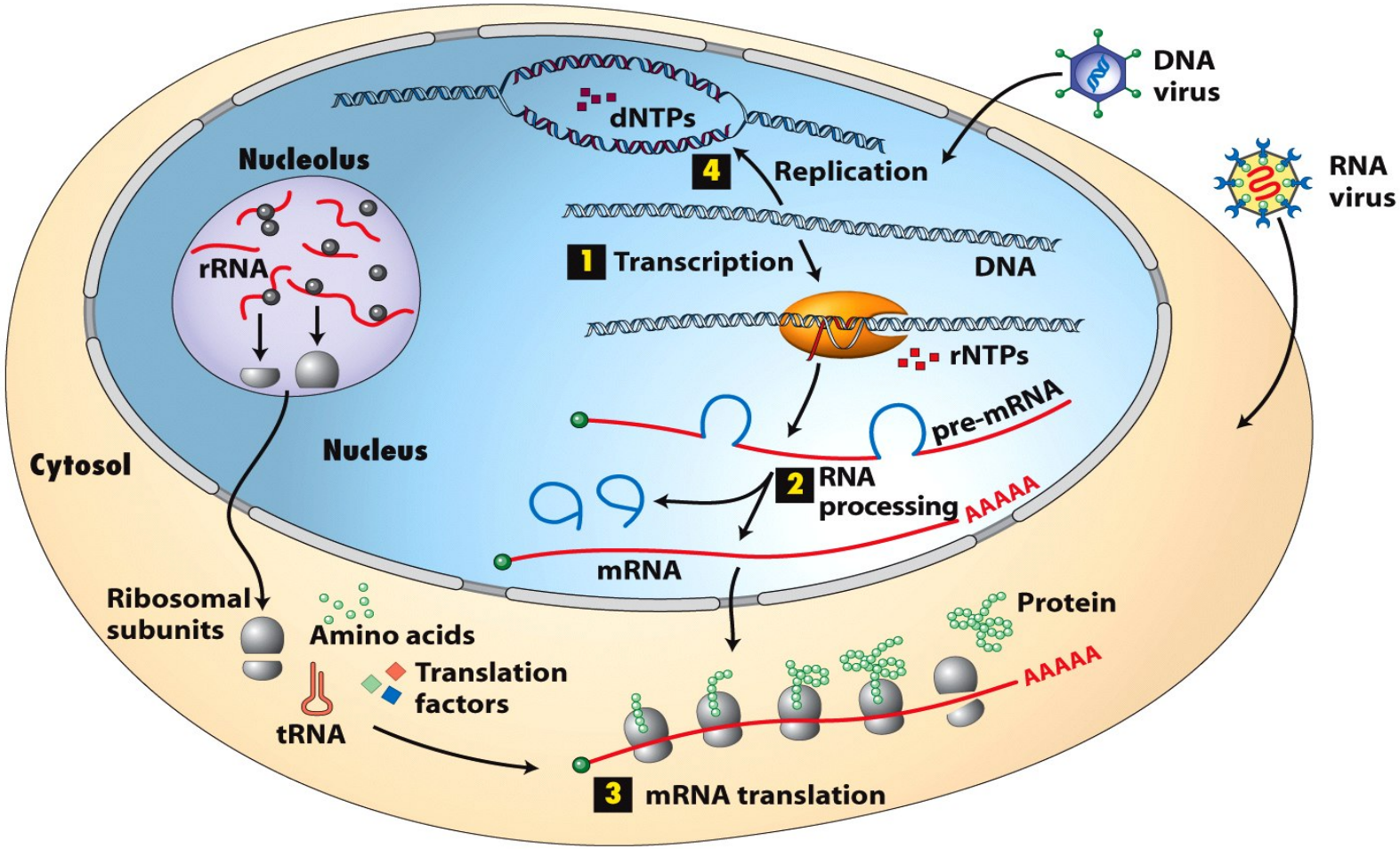
Epigenetic changes influence occurrence of gene expression by turning genes “on” and “off.”
They are triggered by «environmental» changes

**Today we talk about….**

- **The fundaments of Molecular Biology**
  - DNA in the cell
- **The Central Dogma of Molecular Biology**
  - Genetic regulation of Gene expression
- **Epigenetic regulation of Gene expression**
  - DNA Methylation
  - Histone modification
  - Gene Silencing

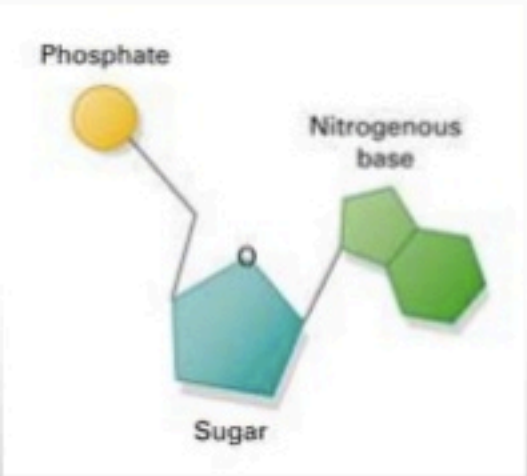
# THE FUNDAMENTS OF MOLECULAR BIOLOGY

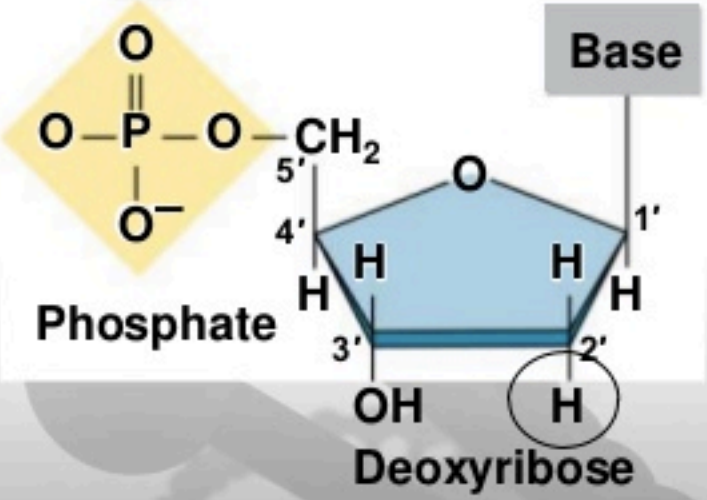
A short journey from the molecules to the sequence




**Molecular biology** is the branch of biology that deals with biological phenomena at the molecular level through the study of DNA and RNA, proteins, and other macromolecules, their <u>biosynthesis</u>, as well as the <u>regulation</u> of their interactions in the various systems of the cell.

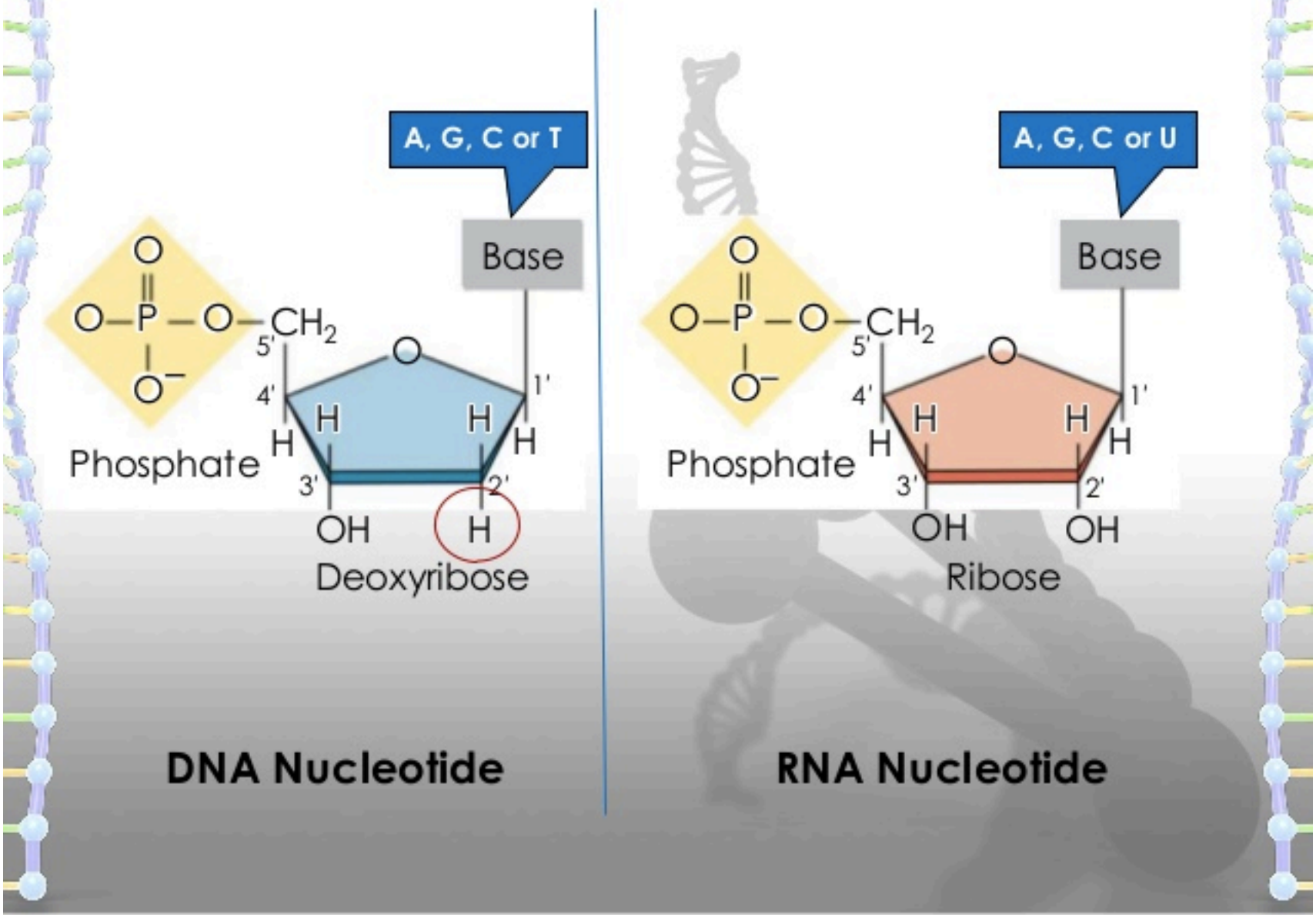
# The famous building block



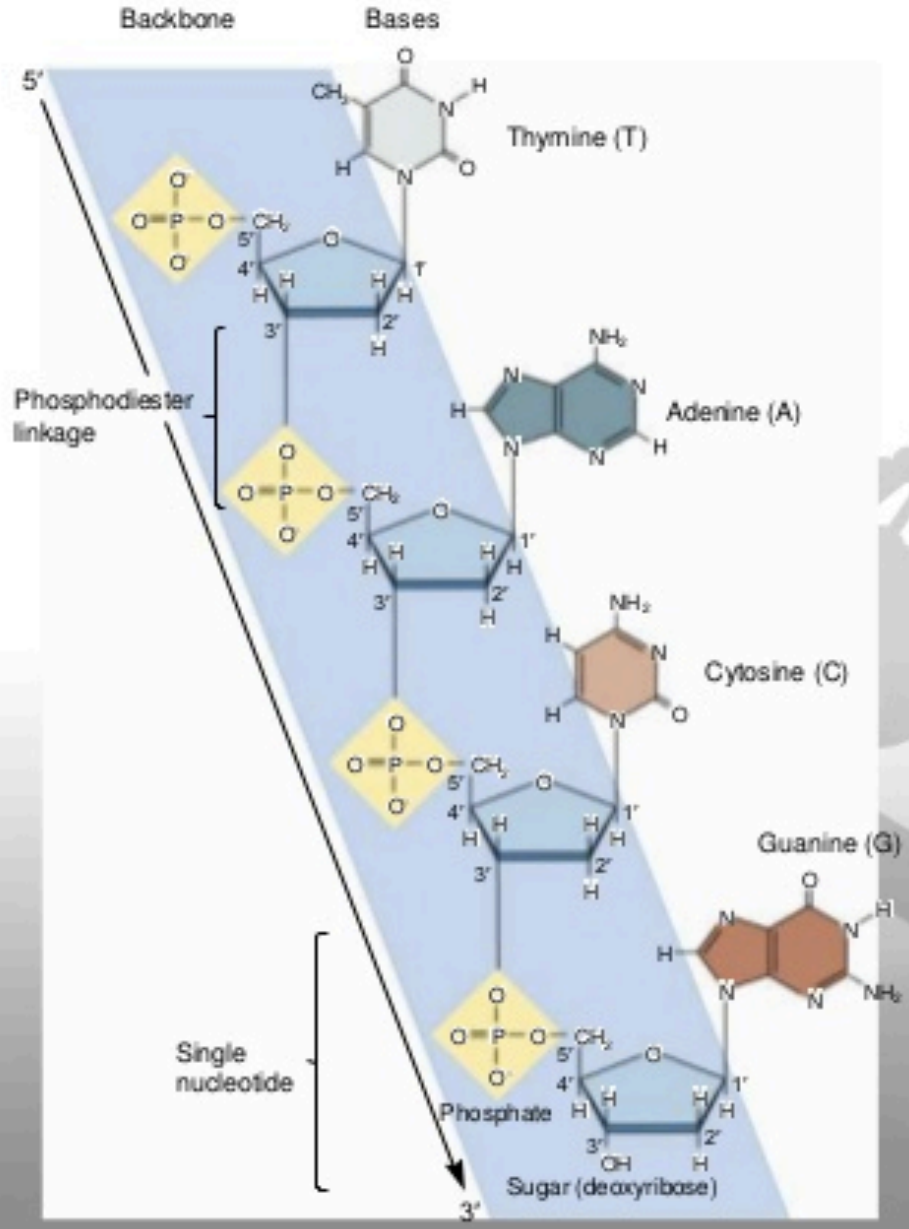




DNA nucleotide

# “different” building blocks for different molecules



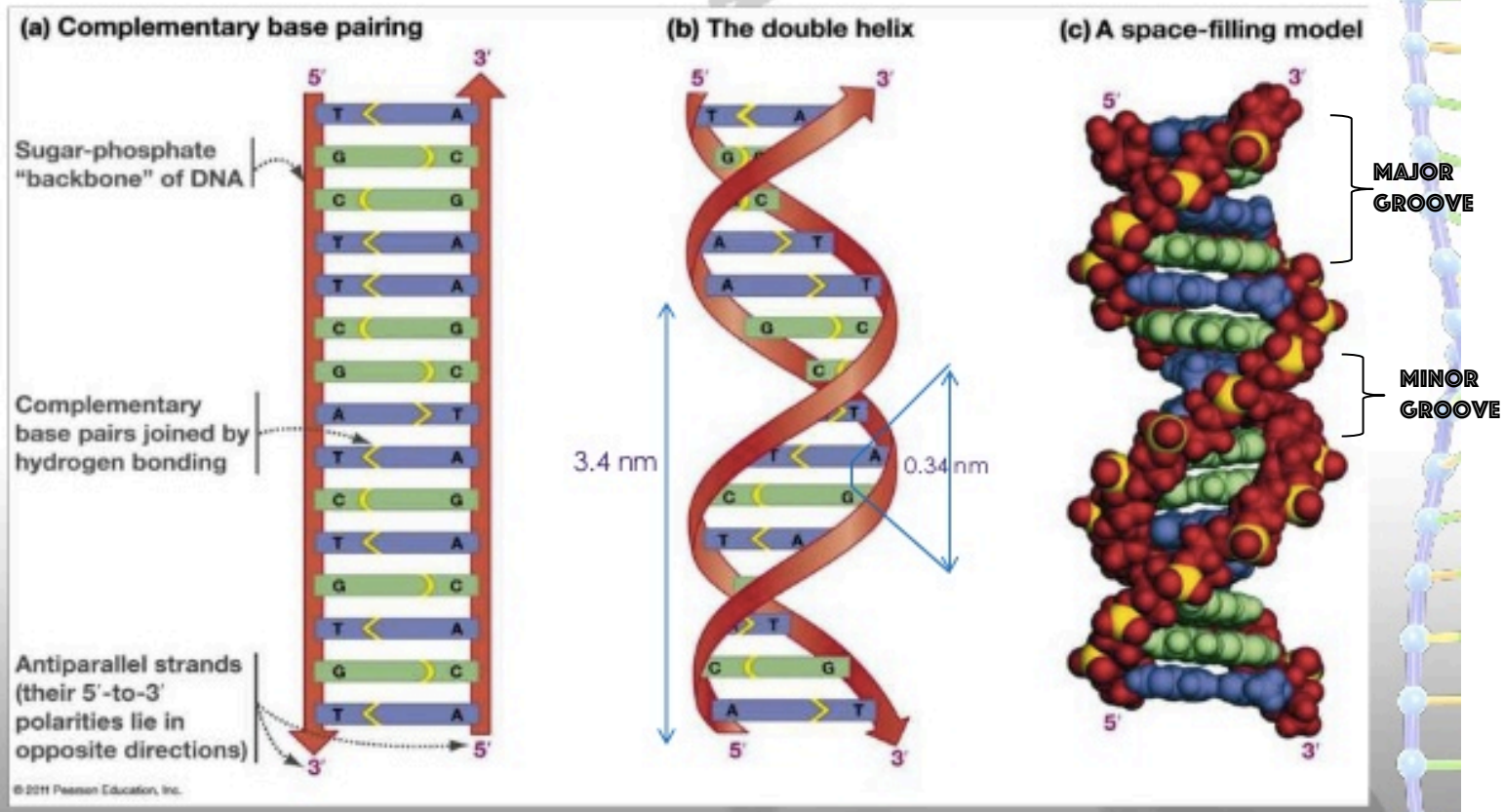
# building blocks in chain



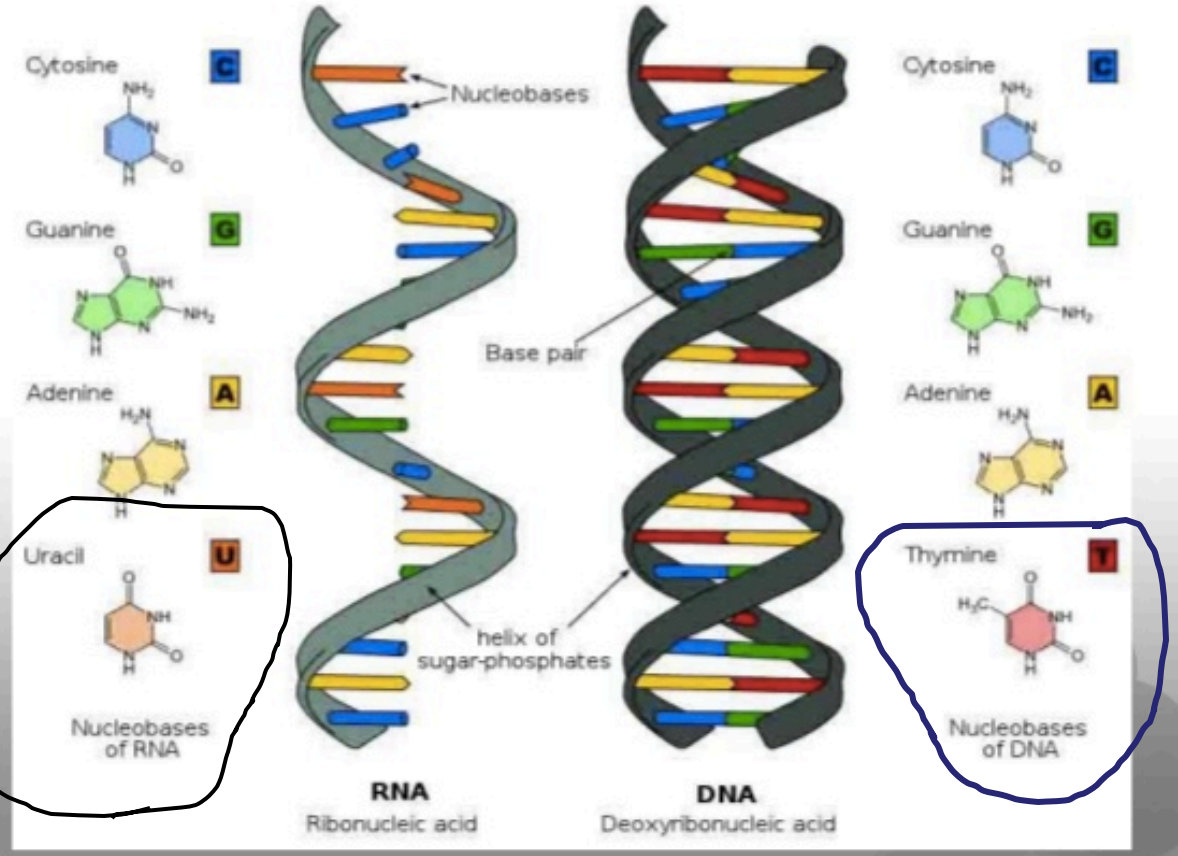
# Question time

```
>gi|528476558|ref|NC_018928.2| Homo sapiens chromosome 17
TCCACTCACTGAGACAATAGACCCAAGCACATCAGCTCTGAGGCCTC
AGCACATCAGCTCTGAGGCCTCCACTCACTGAGACAATAGACCCAAG
TCACTGAGACAATAGACCCAAGCACATCAGCTCTGAGGCCTCCACTC
AACAGCTCTGAGGGCTCCACTCACTGAGACAACAGACCCAAGAACAA
```

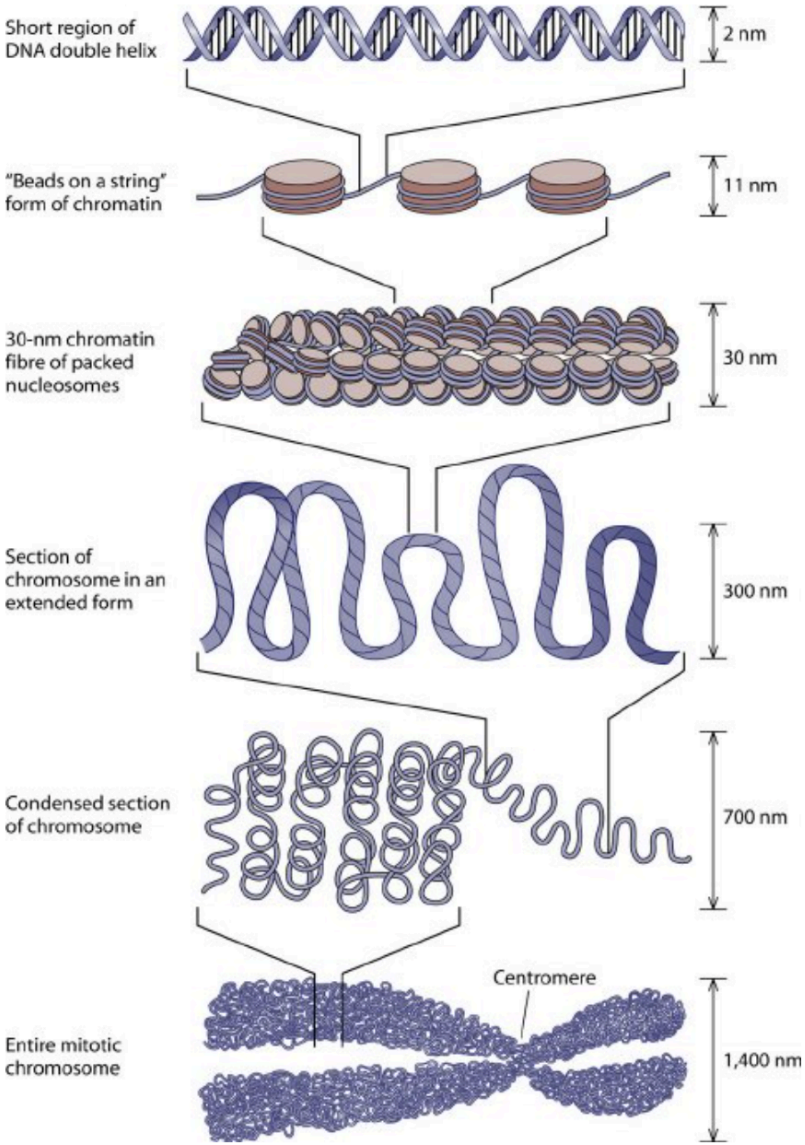
What represent the letters in the FASTA sequence?

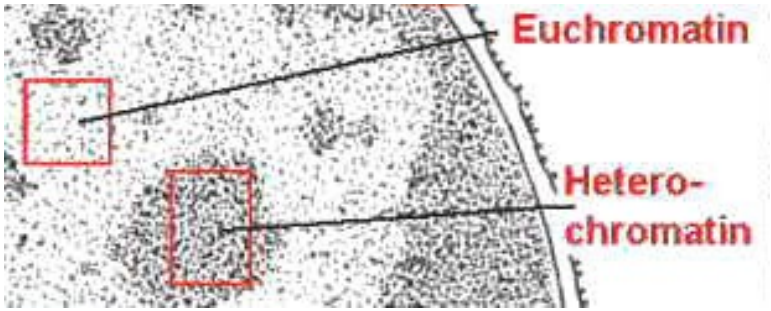

(a) Complementary base pairing
(b) The double helix
(c) A space-filling model
Sugar-phosphate "backbone" of DNA
Complementary base pairs joined by hydrogen bonding
Antiparallel strands (their 5′-to-3′ polarities lie in opposite directions)
5′
3′
3.4 nm
0.34 nm
MAJOR GROOVE
MINOR GROOVE
© 2011 Pearson Education, Inc.

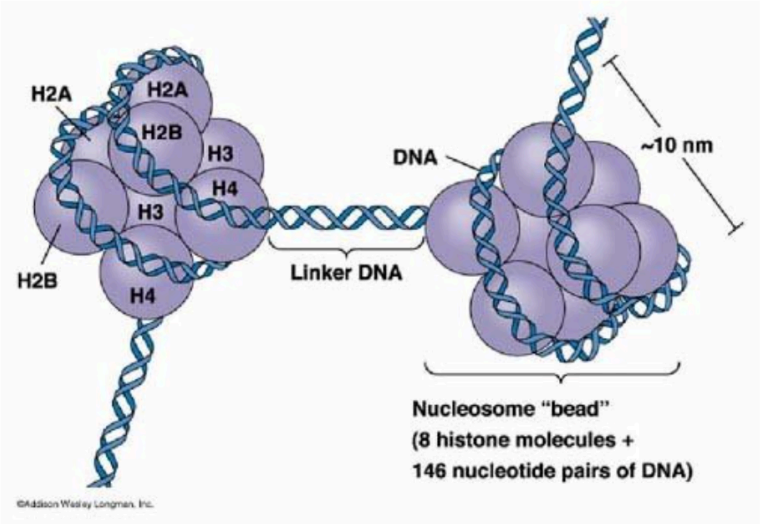
# ❖ RNA V/S DNA


Cytosine
C
Guanine
G
Adenine
A
Uracil
U
Nucleobases of RNA
Nucleobases
Base pair
helix of sugar-phosphates
RNA
Ribonucleic acid
DNA
Deoxyribonucleic acid
Cytosine
C
Guanine
G
Adenine
A
Thymine
T
Nucleobases of DNA

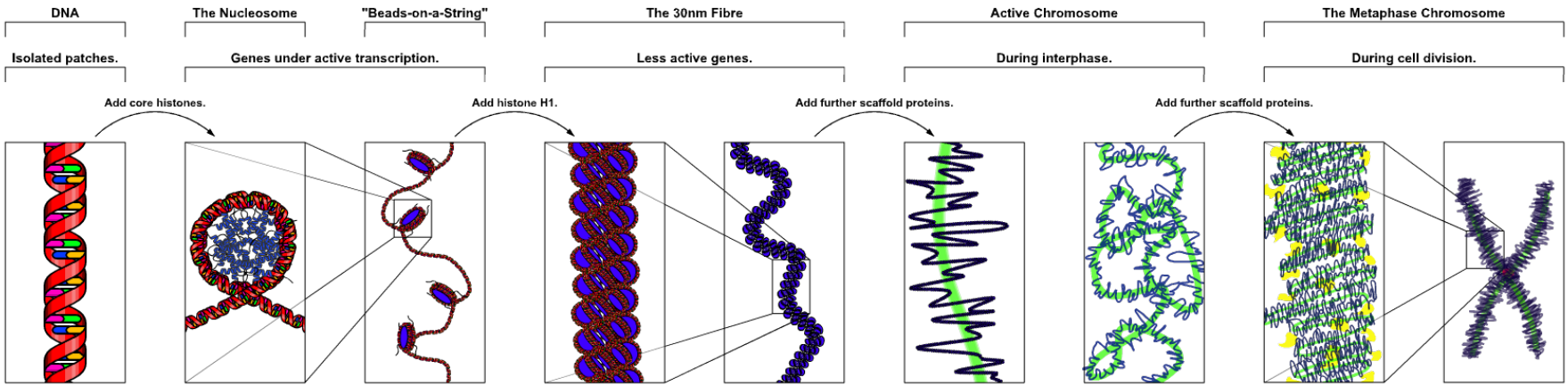
# Chromatin

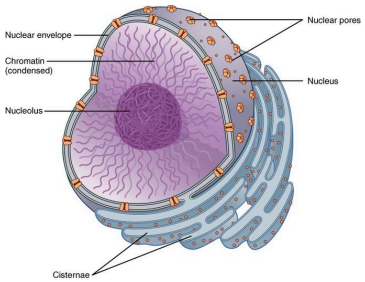

Short region of DNA double helix
2 nm
"Beads on a string" form of chromatin
11 nm
30-nm chromatin fibre of packed nucleosomes
30 nm
Section of chromosome in an extended form
300 nm
Condensed section of chromosome
700 nm
Centromere
Entire mitotic chromosome
1,400 nm


Euchromatin
Hetero-
chromatin


H2A
H2A
H2B
H3
H4
H3
H2B
H4
DNA
~10 nm
Linker DNA
Nucleosome "bead"
(8 histone molecules +
146 nucleotide pairs of DNA)
©Addison Wesley Longman, Inc.

# The chromatine structure changes during cell cycle

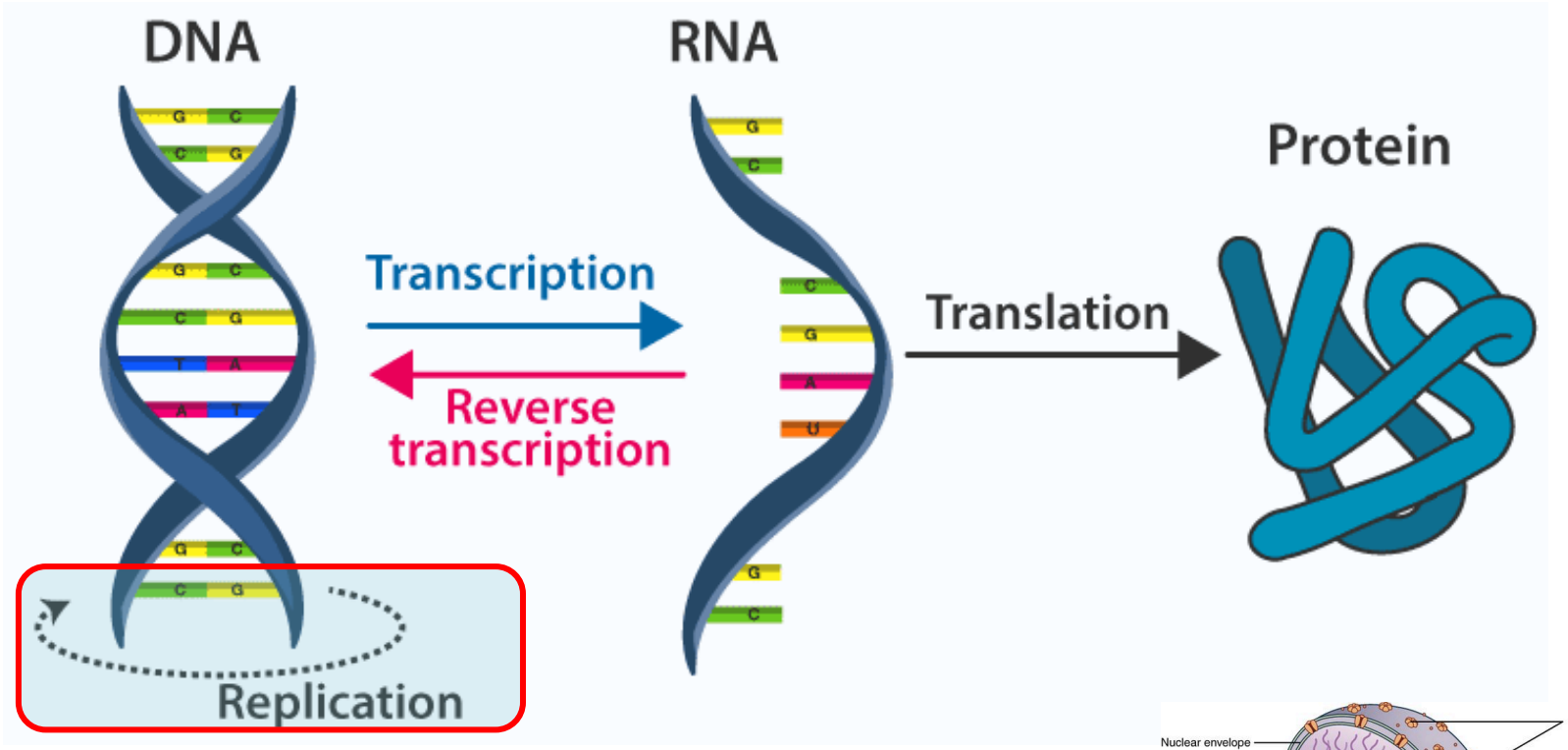


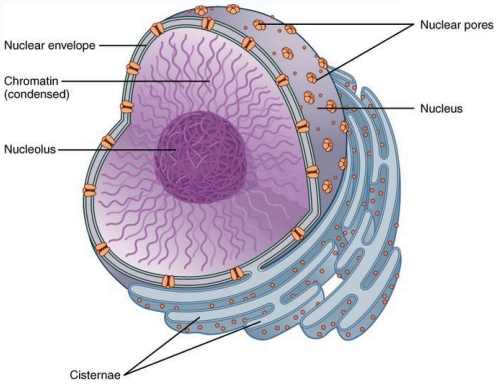



# THE CENTRAL DOGMA OF MOLECULAR BIOLOGY

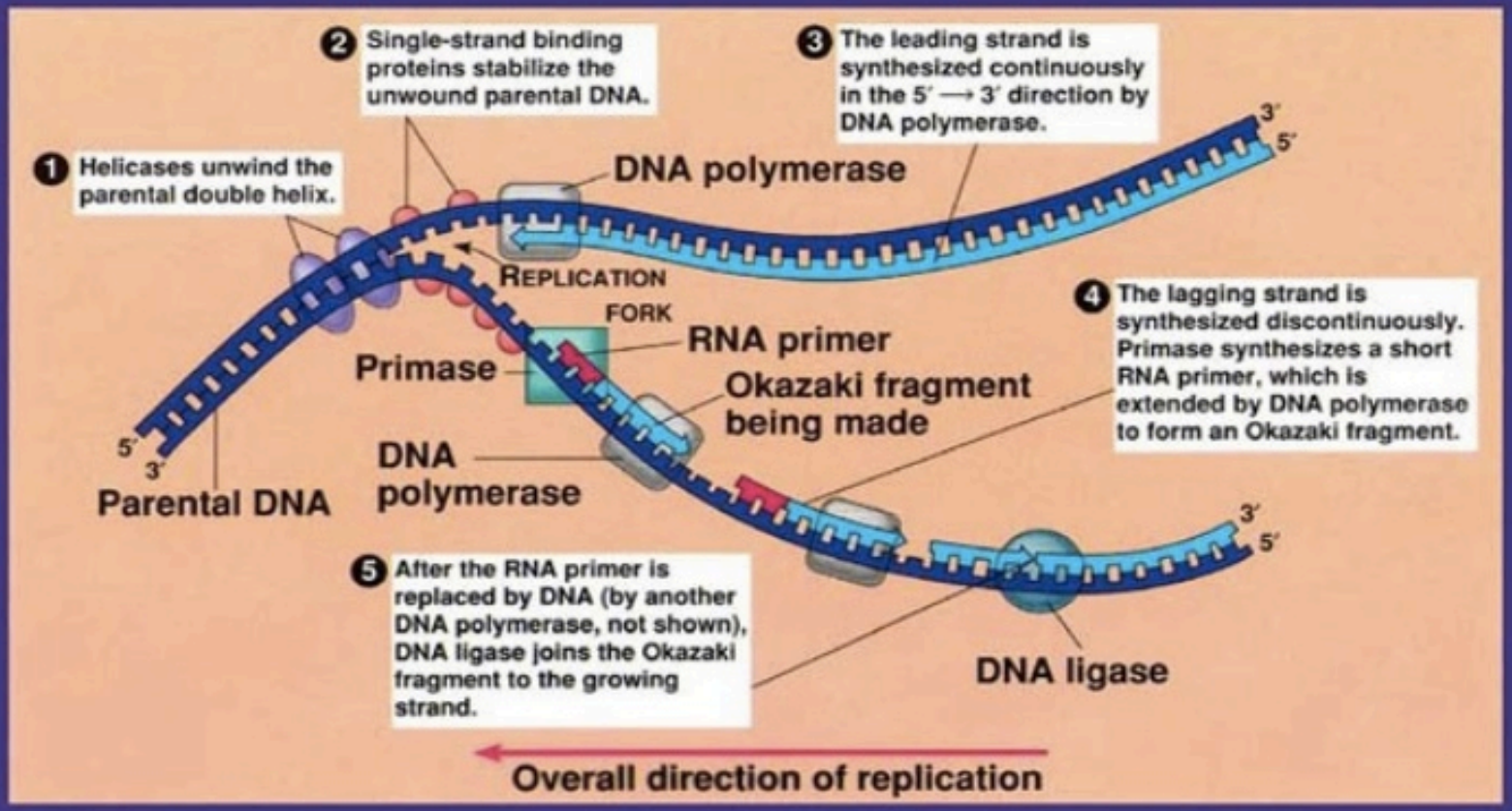
Francis Crick, 1957

The one-way transfer of sequence information from DNA to protein


DNA
RNA
Protein
Transcription
Reverse transcription
Translation
Replication


Nuclear envelope
Chromatin (condensed)
Nucleolus
Nuclear pores
Nucleus
Cisternae

# A SUMMARY OF DNA REPLICATION

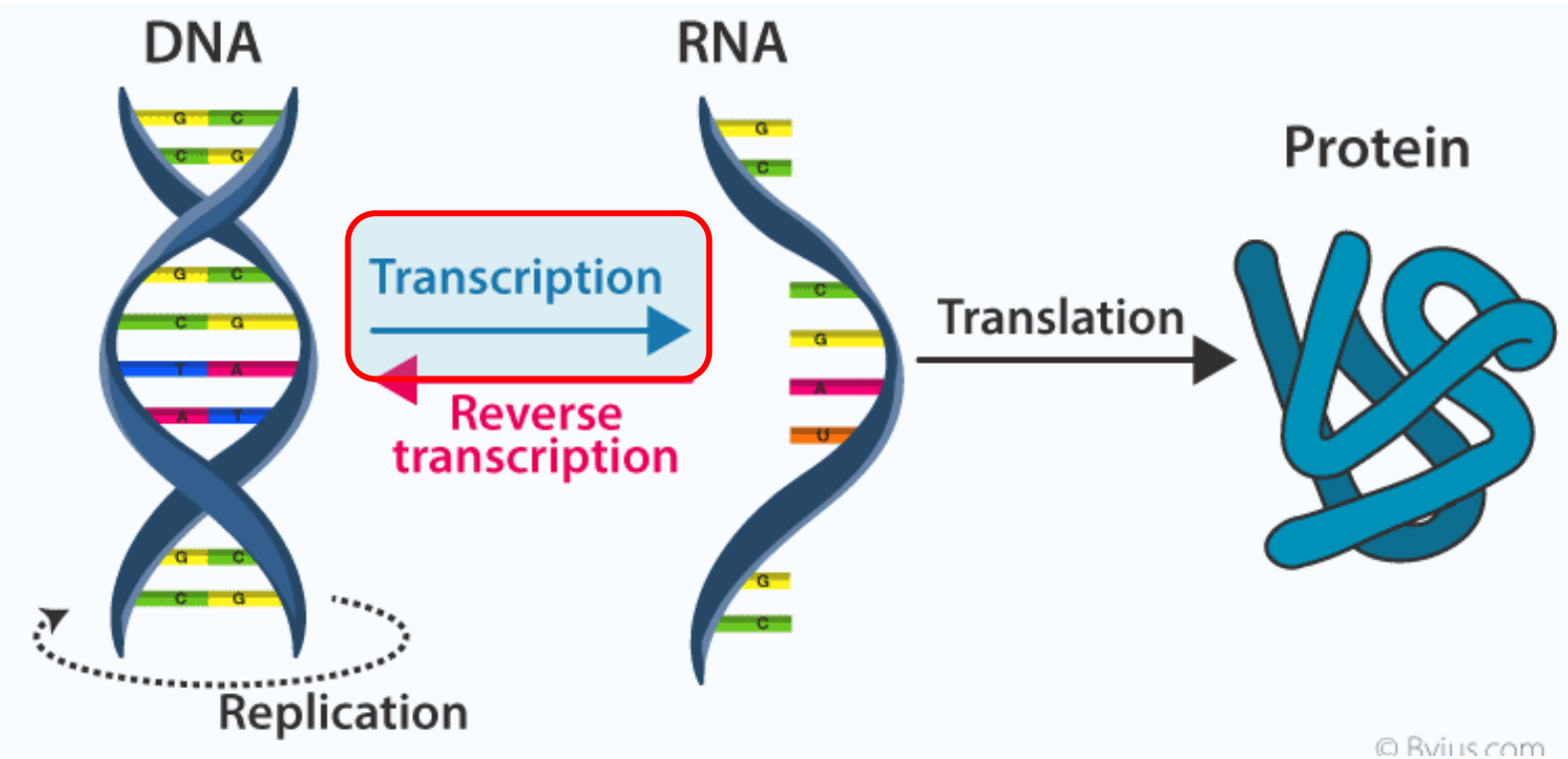


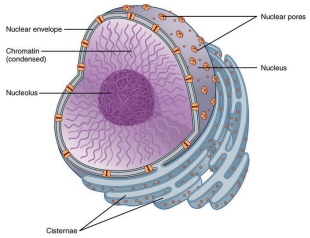
# Question time

```
>gi|528476558|ref|NC_018928.2| Homo sapiens chromosome 17
TCCACTCACTGAGACAATAGACCCAAGCACATCAGCTCTGAGGCCTC
AGCACATCAGCTCTGAGGCCTCCACTCACTGAGACAATAGACCCAAG
TCACTGAGACAATAGACCCAAGCACATCAGCTCTGAGGCCTCCACTC
AACAGCTCTGAGGGCTCCACTCACTGAGACAACAGACCCAAGAACAA
```

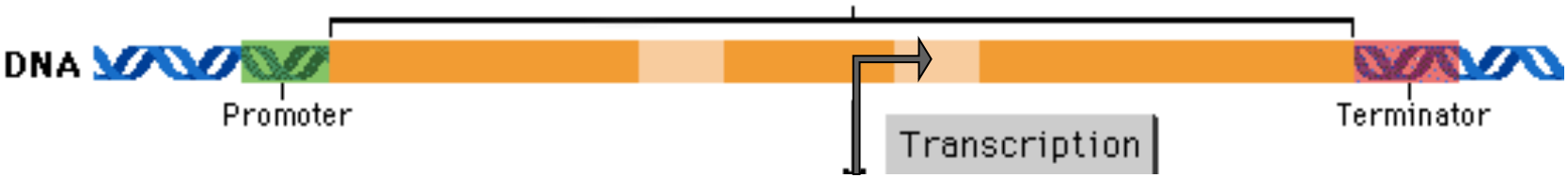
Which strand is represented and what is the sequence direction?

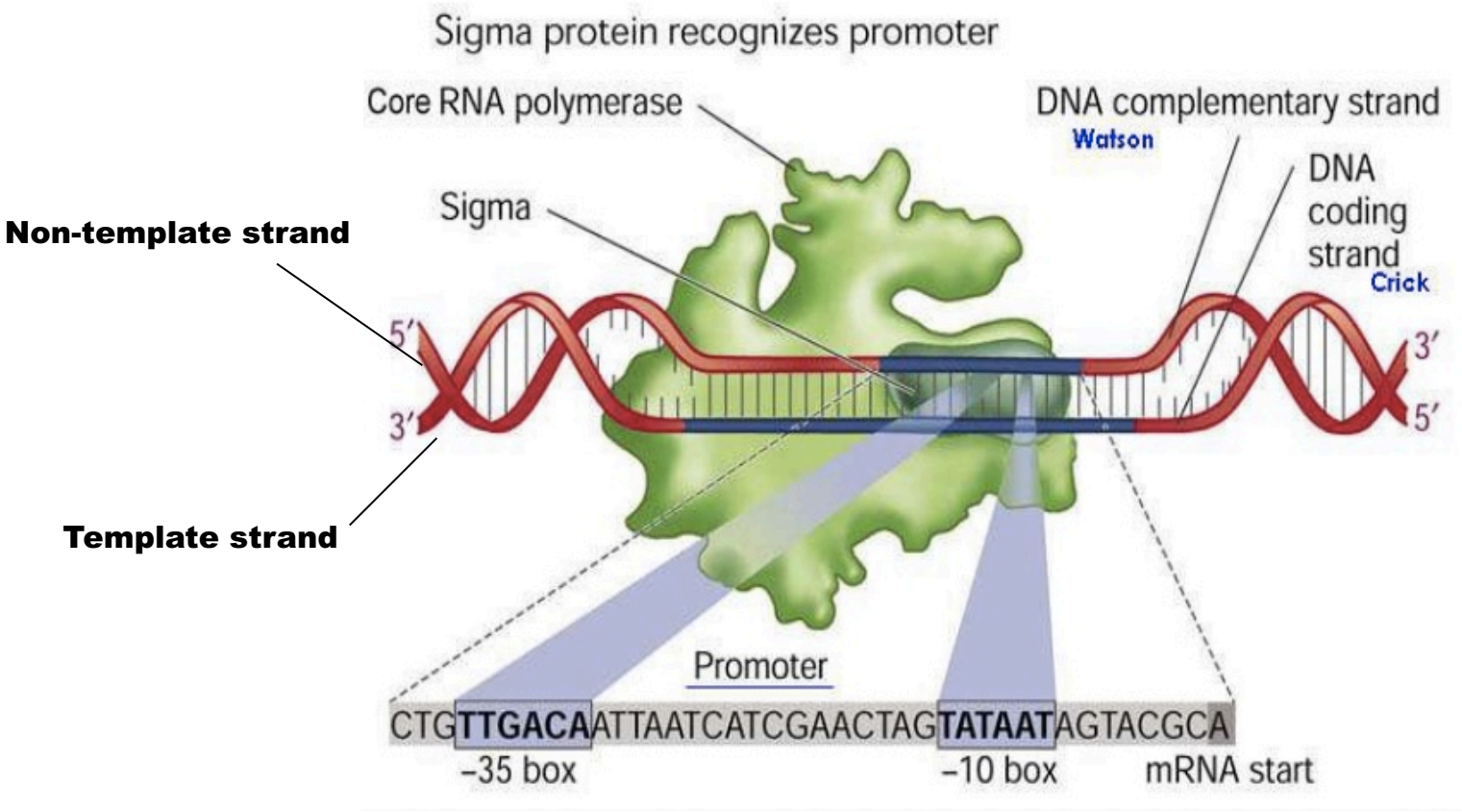
# The Central dogma of Molecular Biology



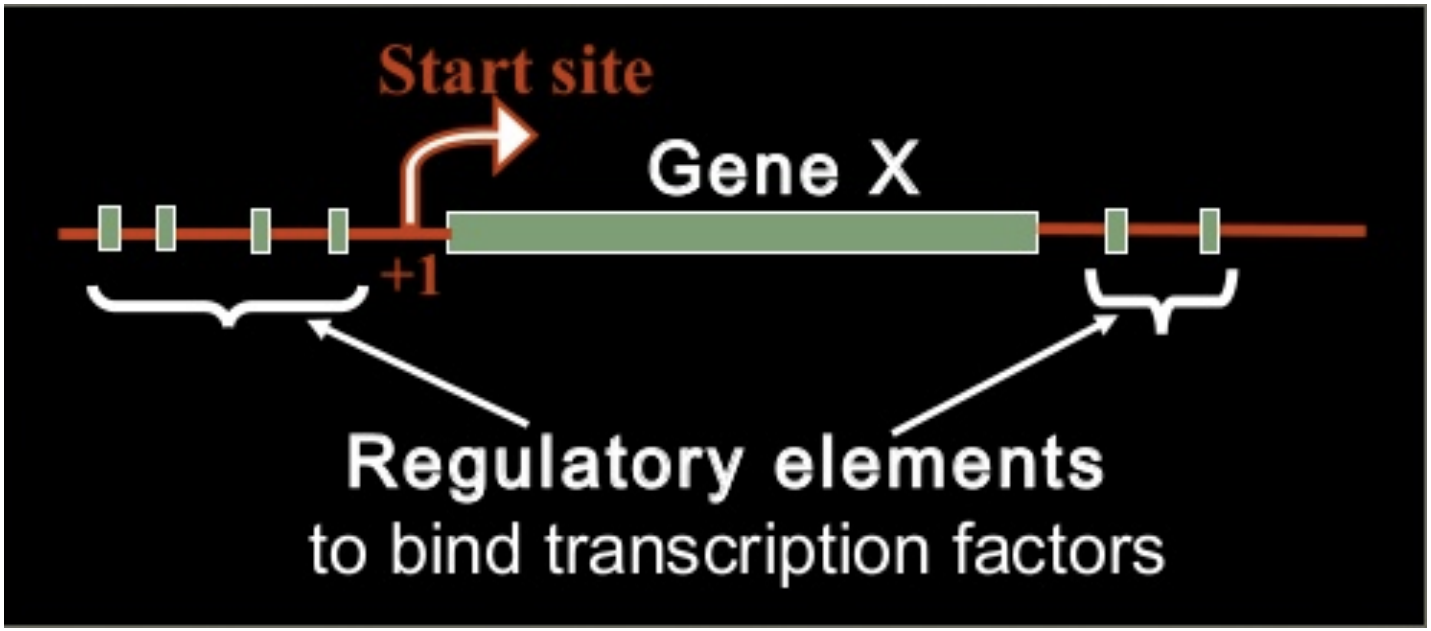



# The gene


DNA
Promoter
Transcription
Terminator


Sigma protein recognizes promoter
Core RNA polymerase
DNA complementary strand
Watson
DNA coding strand
Crick
Sigma
Non-template strand
Template strand
5′
3′
3′
5′
Promoter
CTGTTGACAATTAATCATCGAACTAGTATAATAGTACGCA
–35 box
–10 box
mRNA start

# Regulation of Gene expression: **Transcription factors**



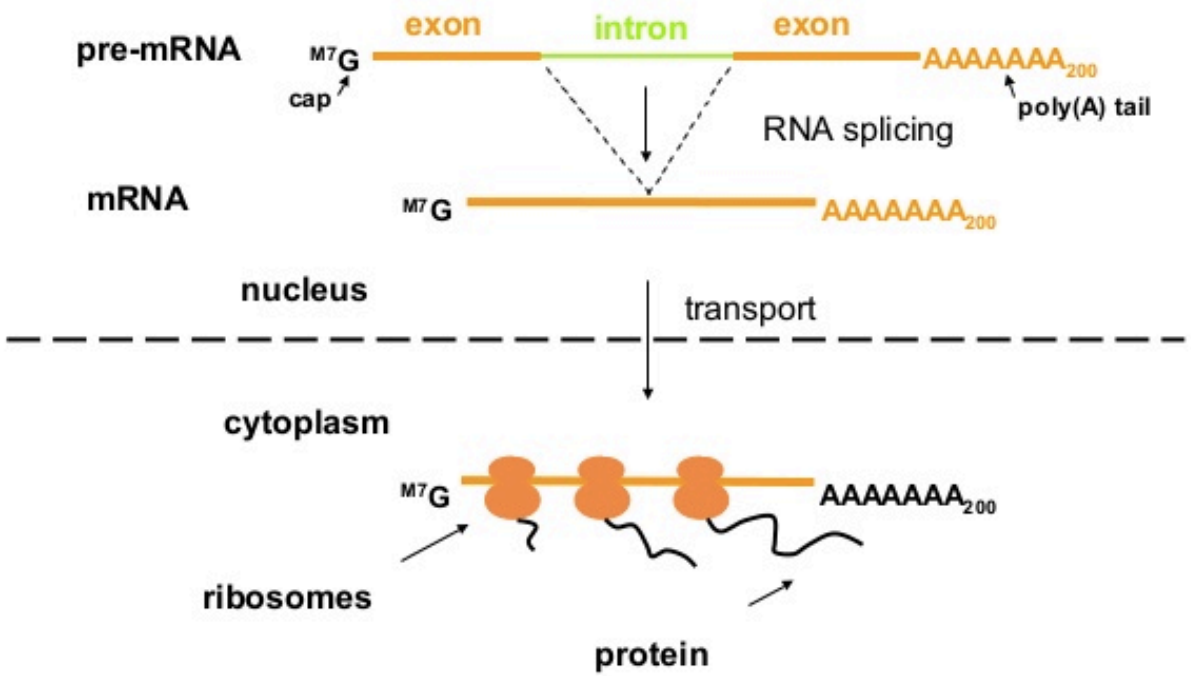

- TFs are proteins that binds to specific sequence of DNA(*cis*-acting)elements in promoter region and regulate transcription

- Number of transcription factors depends on genome size

- Use different mechanism to up or down regulate gene expression

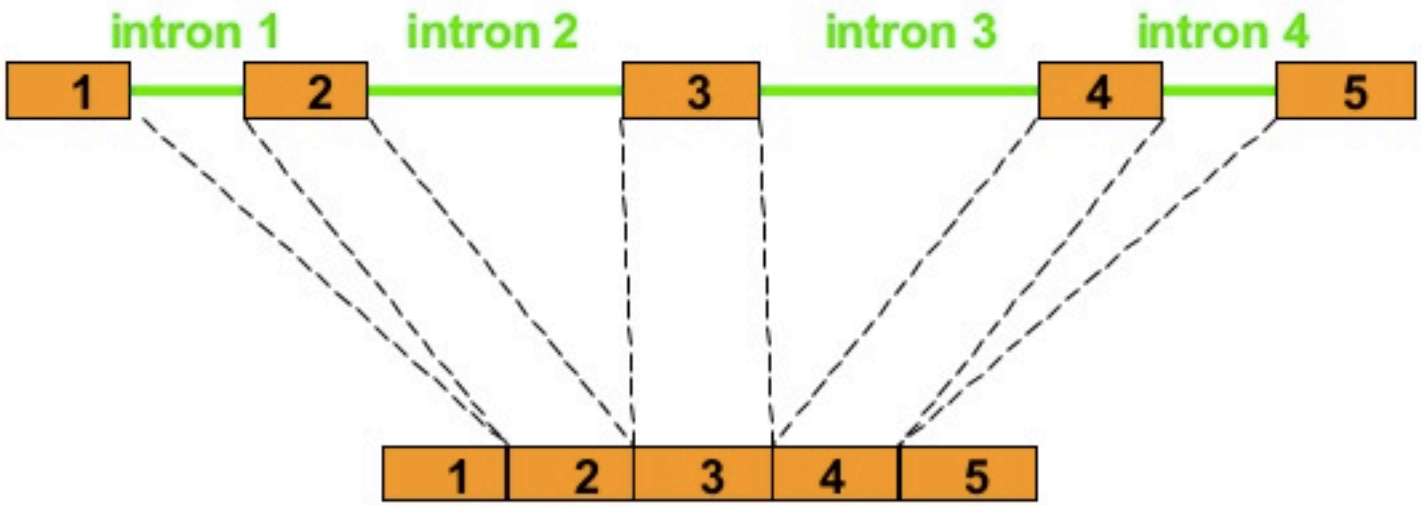
# Regulation of Gene expression: **Splicing**

## Pre-messenger RNA Processing



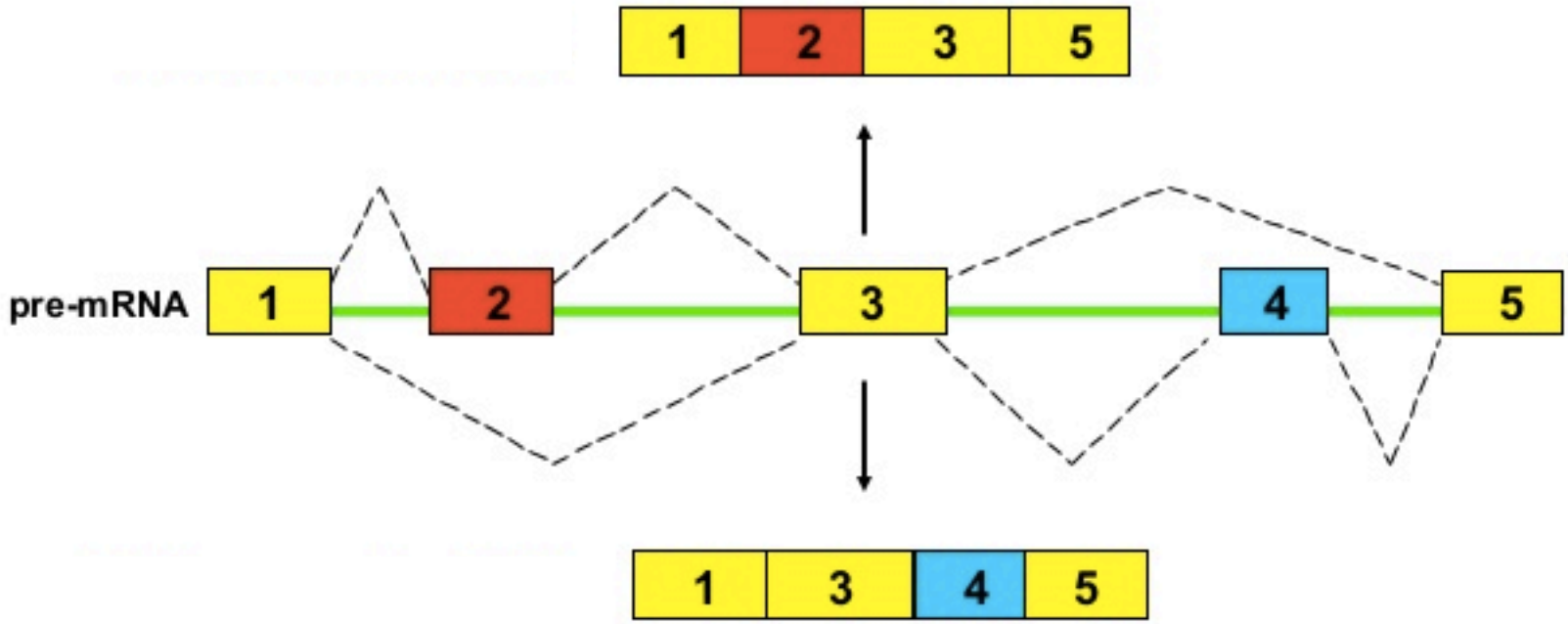
# Alternative splicing

Can generate tens of thousands of mRNAs from a single primary transcript




Usually all introns must be removed before the mRNA can be translated to produce protein

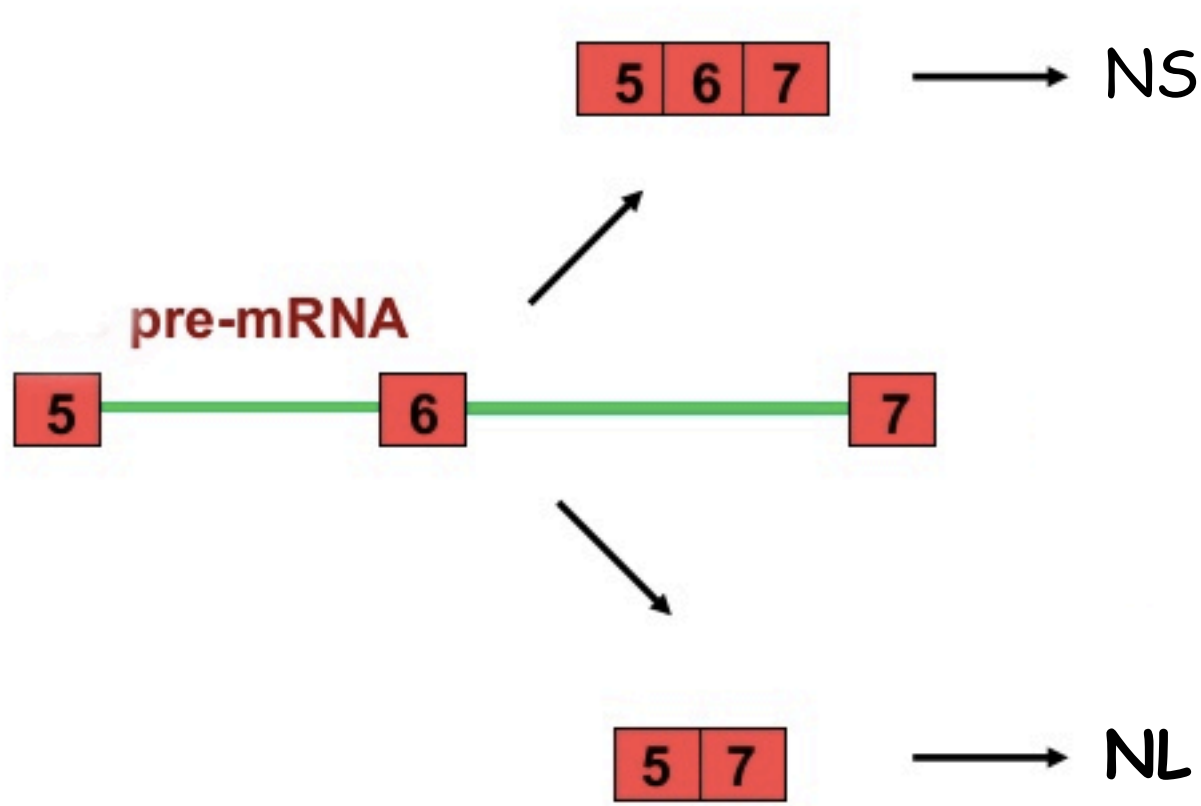
However, multiple introns may be spliced differently in different circumstances, for example in different tissues.




Thus one gene can encode more than one protein. The proteins are similar but not identical and may have distinct properties. This is important in complex organisms

## Alternative splicing can generate mRNAs encoding proteins with different, even opposite functions

**The N resistance gene**



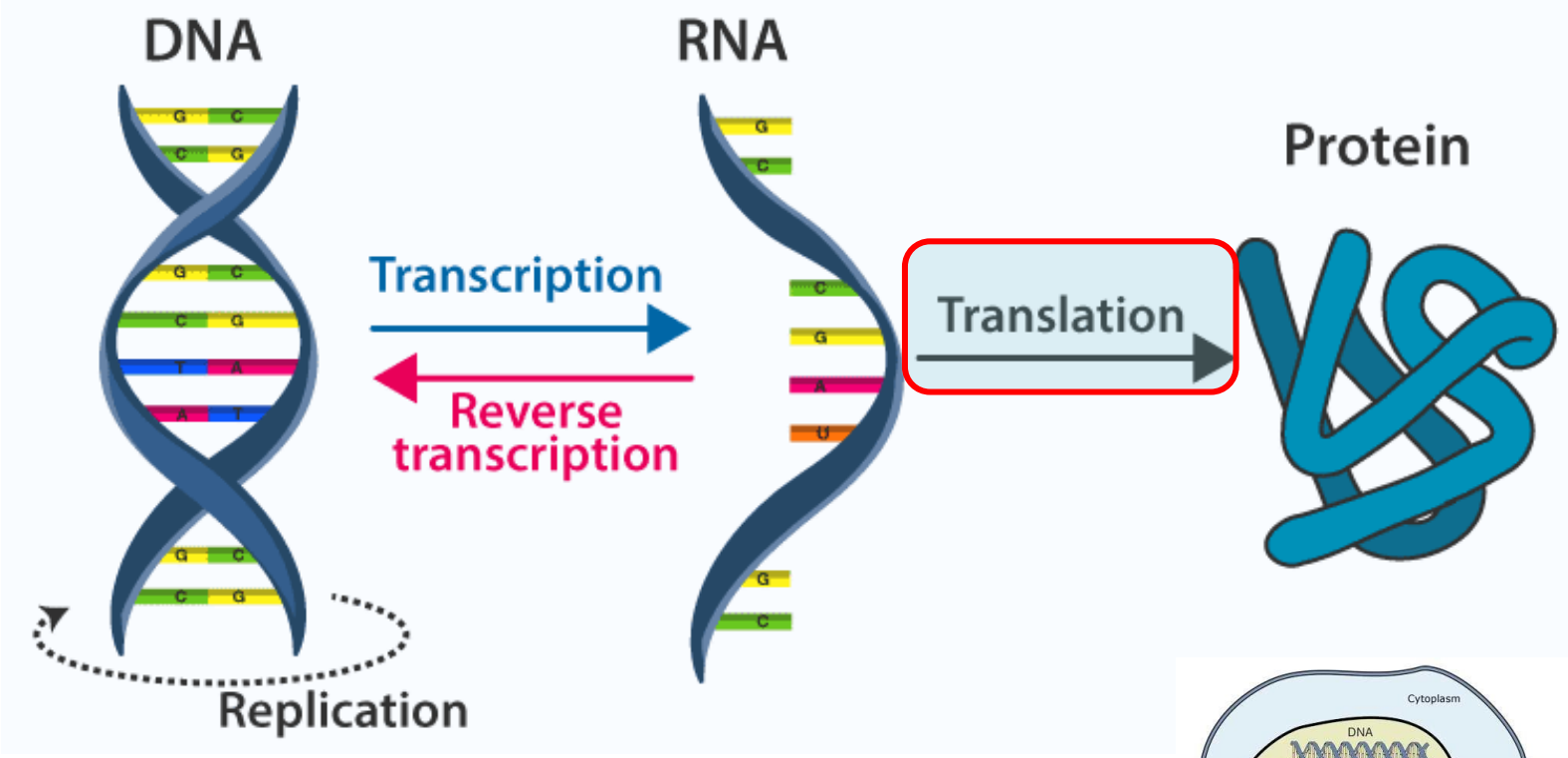
**Resistance to TMV infection**

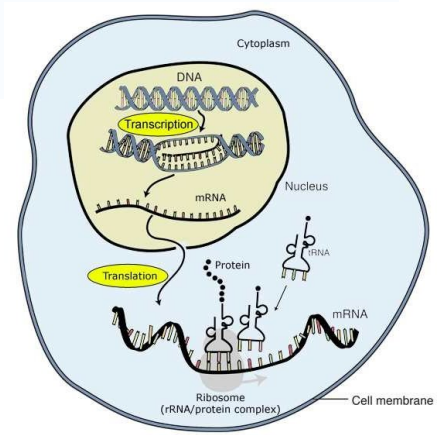
## Question time

Which parts of the pre-mRNA are included in a CDS sequence?

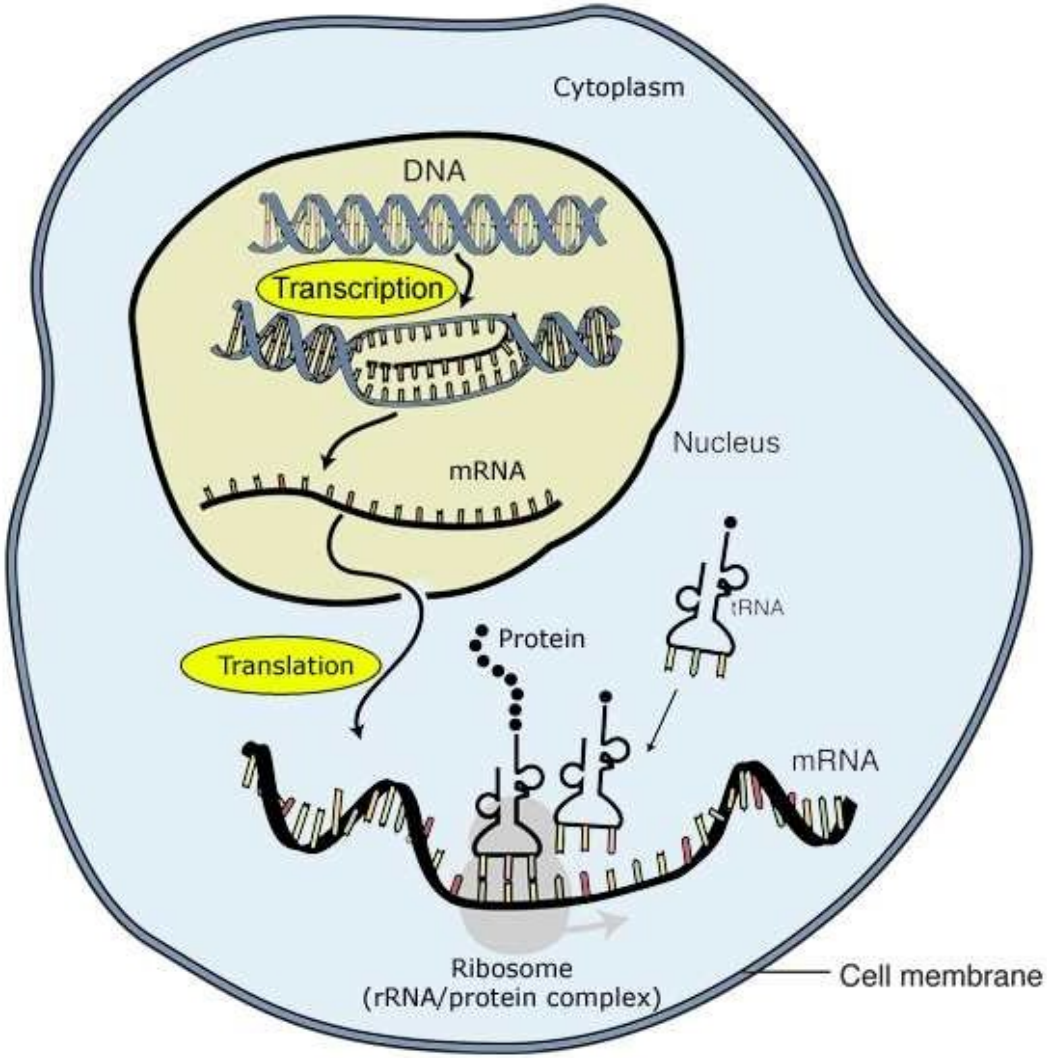
CDS stands for coding sequence

# The Central dogma of Molecular Biology

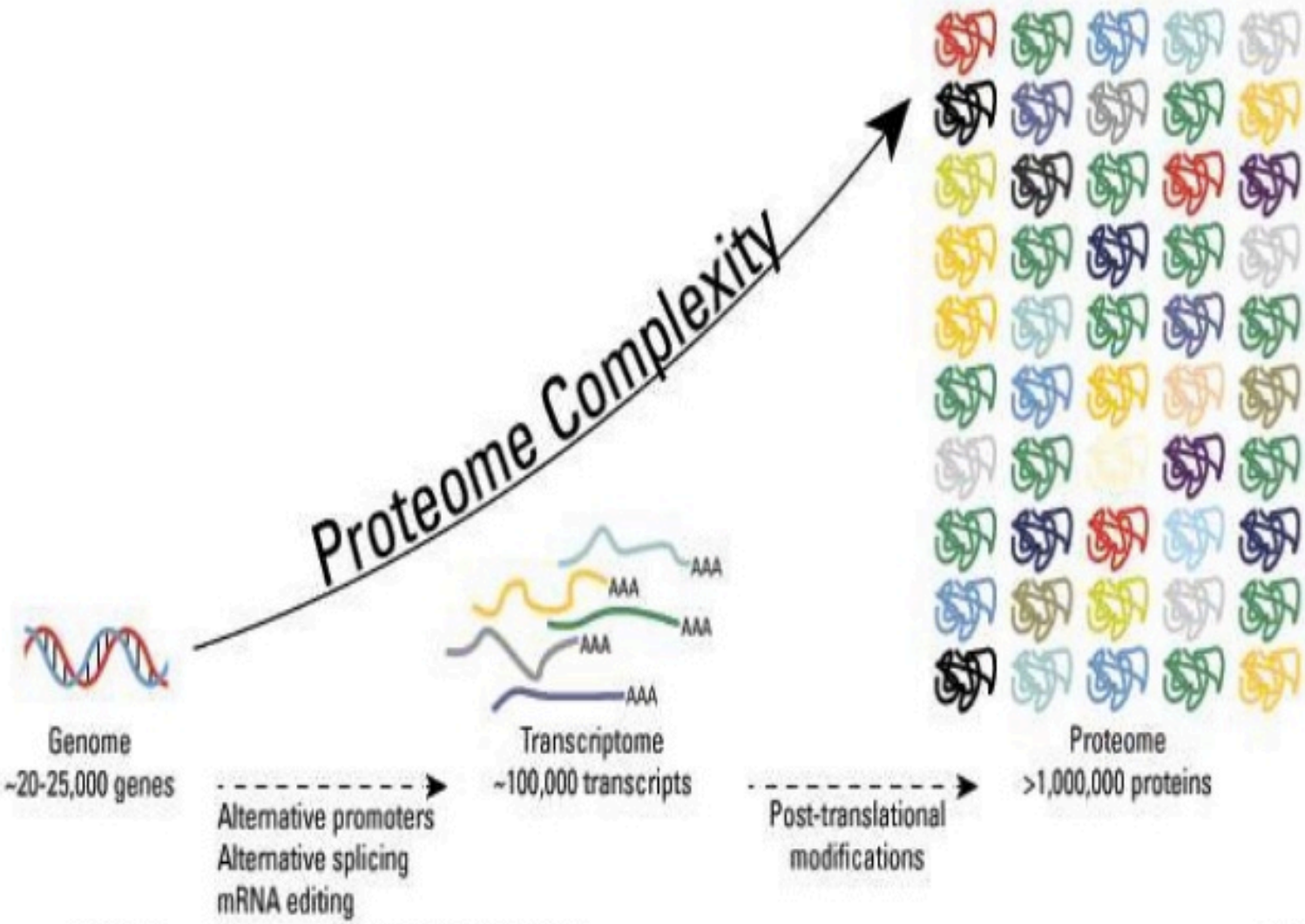


# The players of mRNA translation:




Messenger, mRNA
Ribosomes
Transfer, tRNA
Amino Acids


Proteome Complexity
AAA
AAA
AAA
AAA
Genome
~20-25,000 genes
Alternative promoters
Alternative splicing
mRNA editing
Transcriptome
~100,000 transcripts
Post-translational
modifications
Proteome
>1,000,000 proteins

Gene silencing

# EPIGENETIC REGULATION OF GENE EXPRESSION

consists of

Potentially heritable changes in the **phenotype** without a change in the underlining **genotype**.

can be mediated by

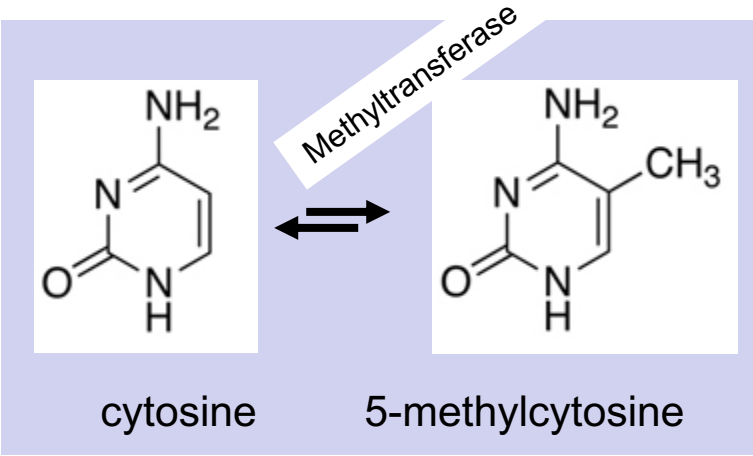
**DNA METHYLATION**

**HISTONE MODIFICATION**

**NON-CODING RNA**

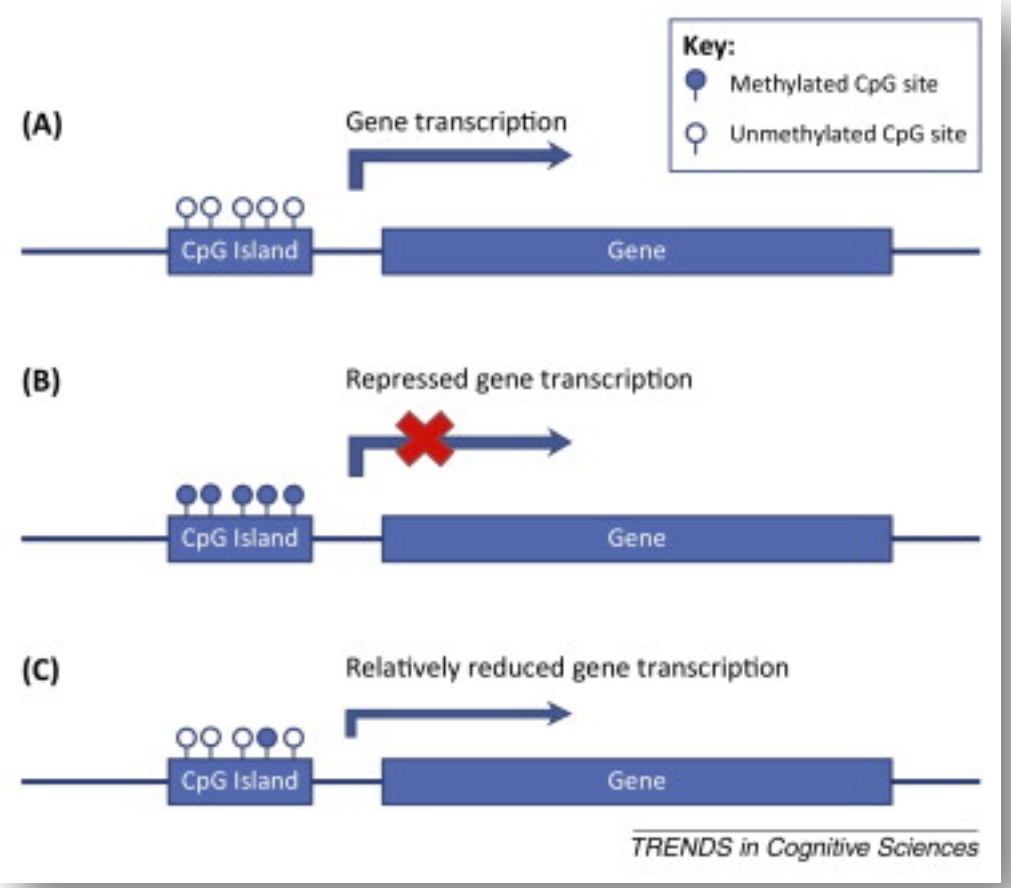
# DNA METHYLATION

DNA can be covalently modified by cytosine methylation.

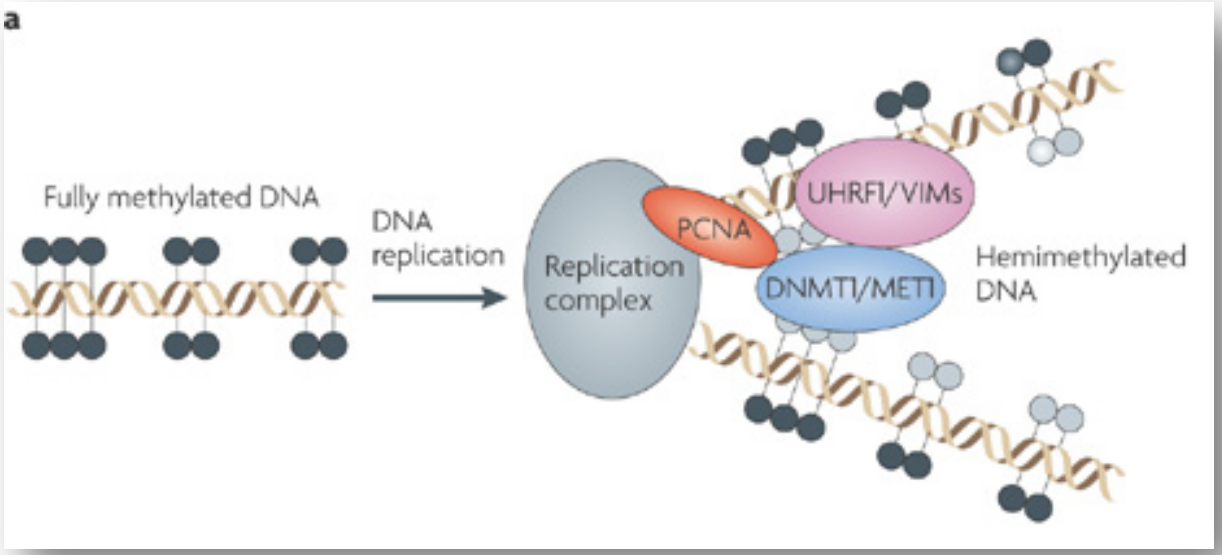


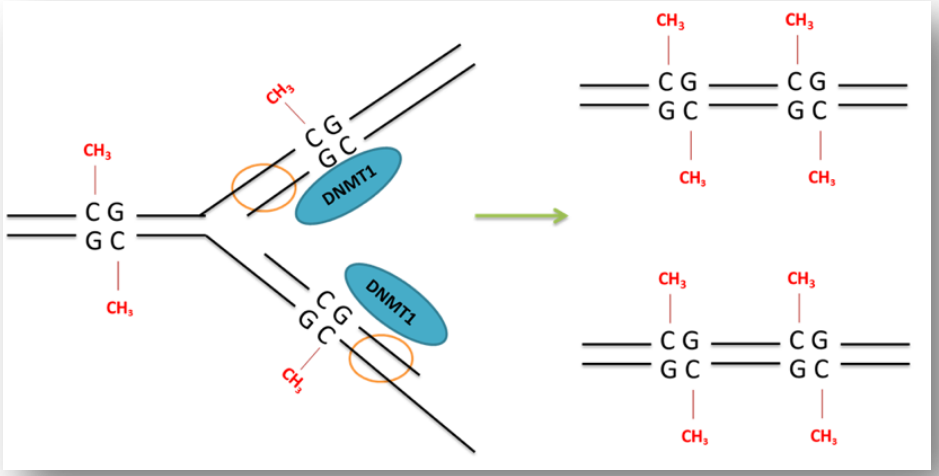
# DNA methylation

Methylated regions in the promoter inhibit/reduce the docking of the transcription complex

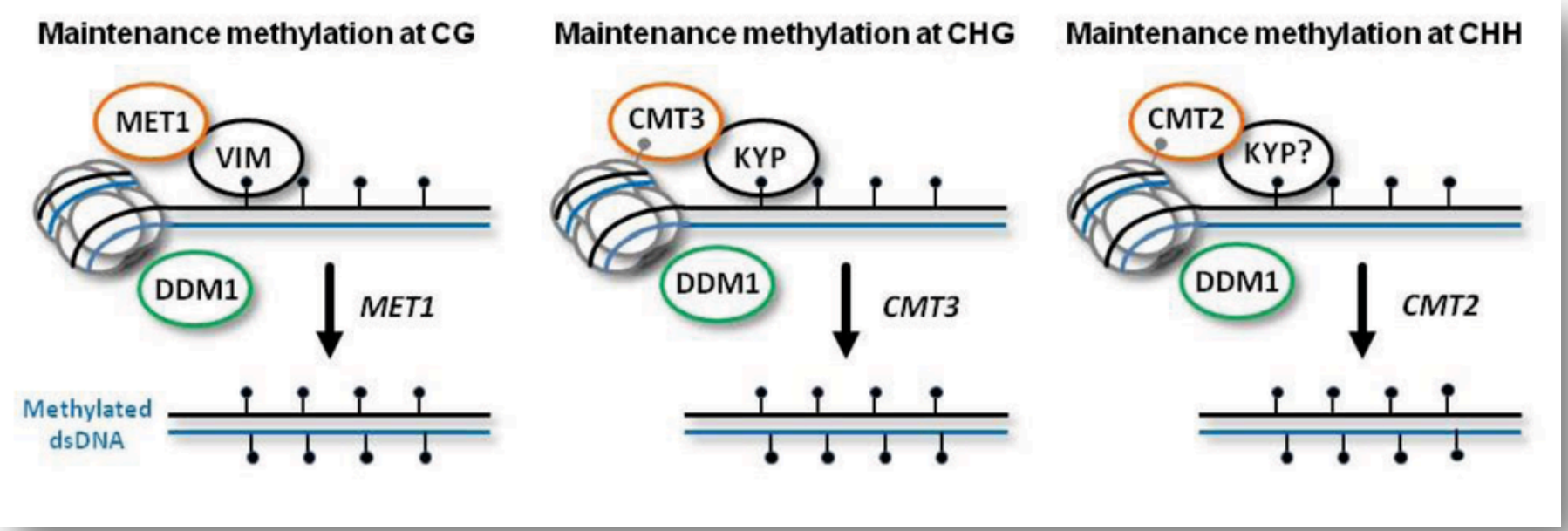


# Maintenance of DNA methylation



# Maintenance of DNA methylation



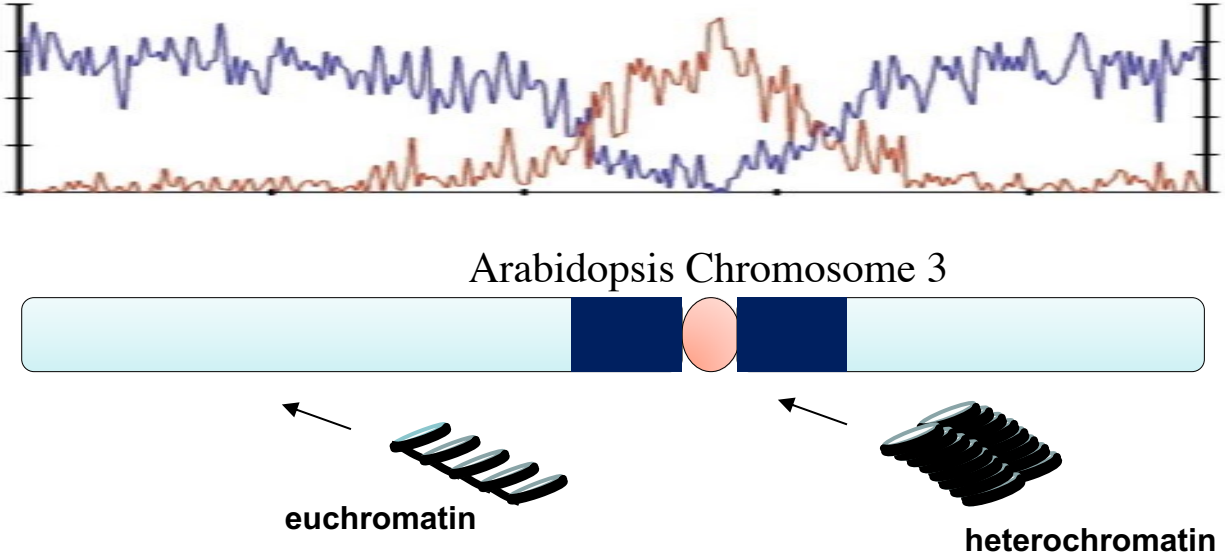

In plants there are 3 methylation sites: **CG** **CHG** **CHH**

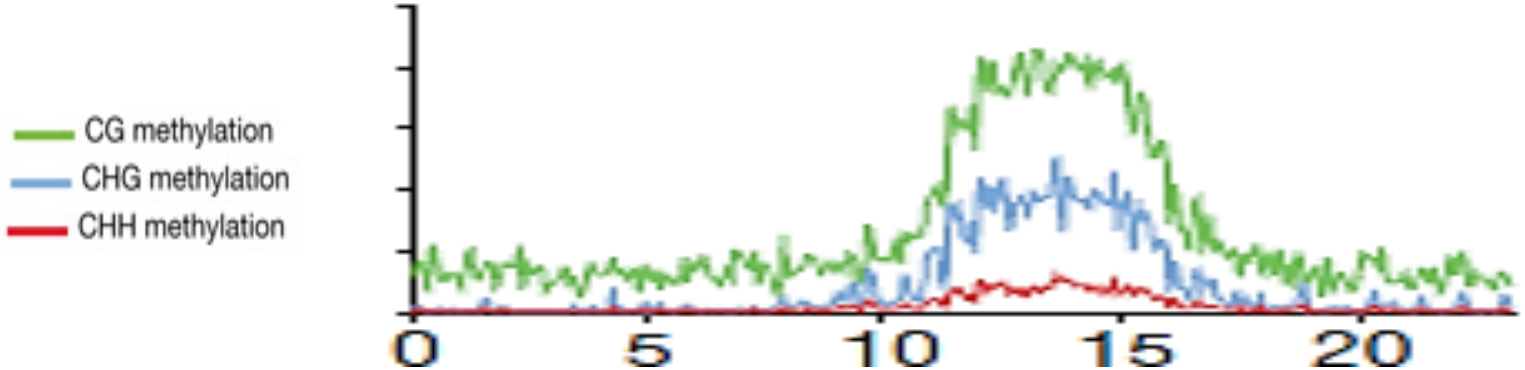
H= A, T, C

# Distribution of methylated regions

**Gene density**

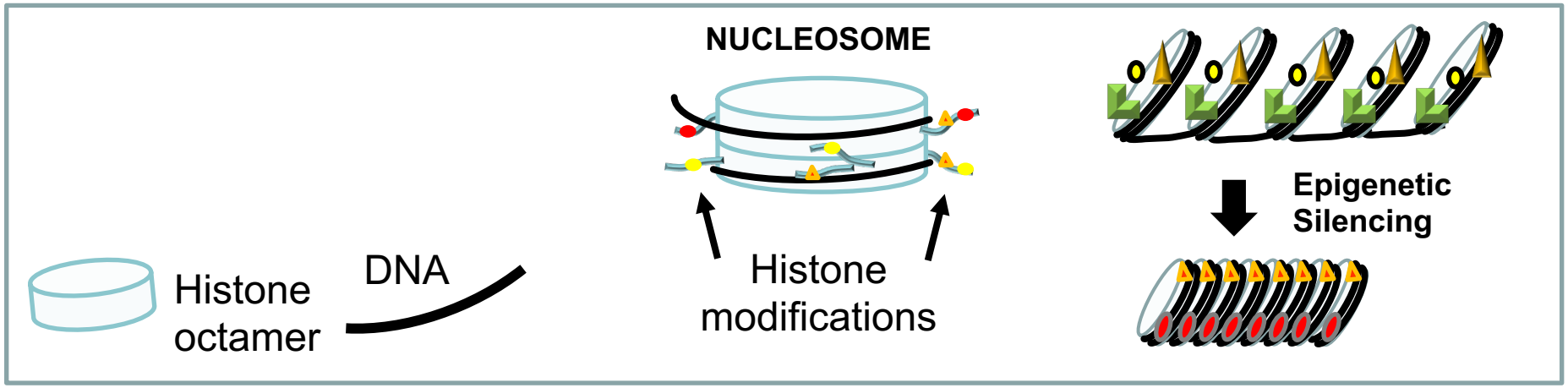
**Repetitive element and transposon density**



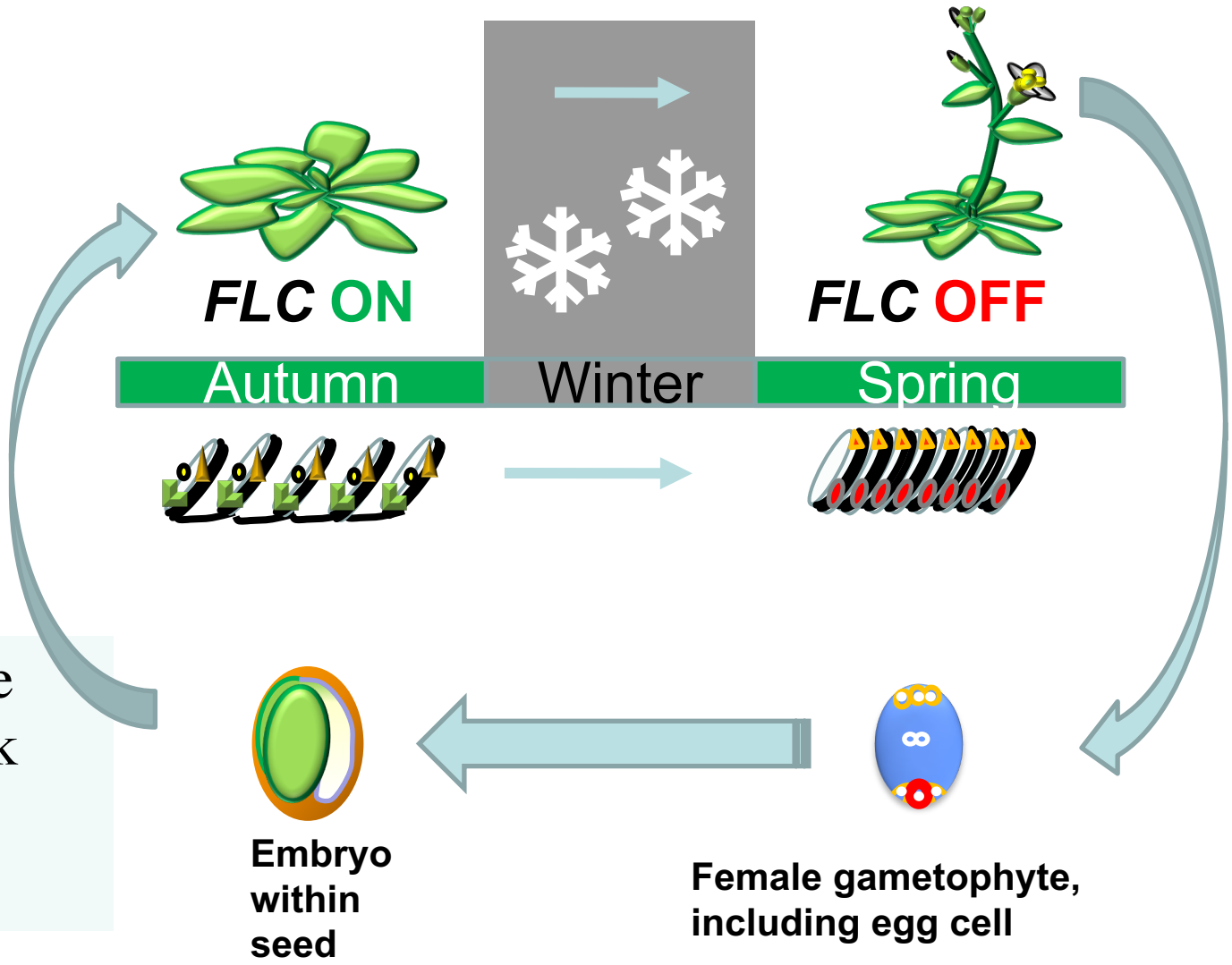
# HISTONE MODIFICATION

Histone proteins of the nucleosome can be covalently modified, affecting chromatin structure and disrupting or facilitating chromatin-associated proteins



# Epigenetic regulation of FLC (Flowering Locus C) in Arabidopsis

*FLC* inhibits flowering




Resetting of the epigenetic mark during reproduction

# NON-CODING RNA

*the majority of the genome is transcribed into RNA transcripts, most of which do not code for protein*

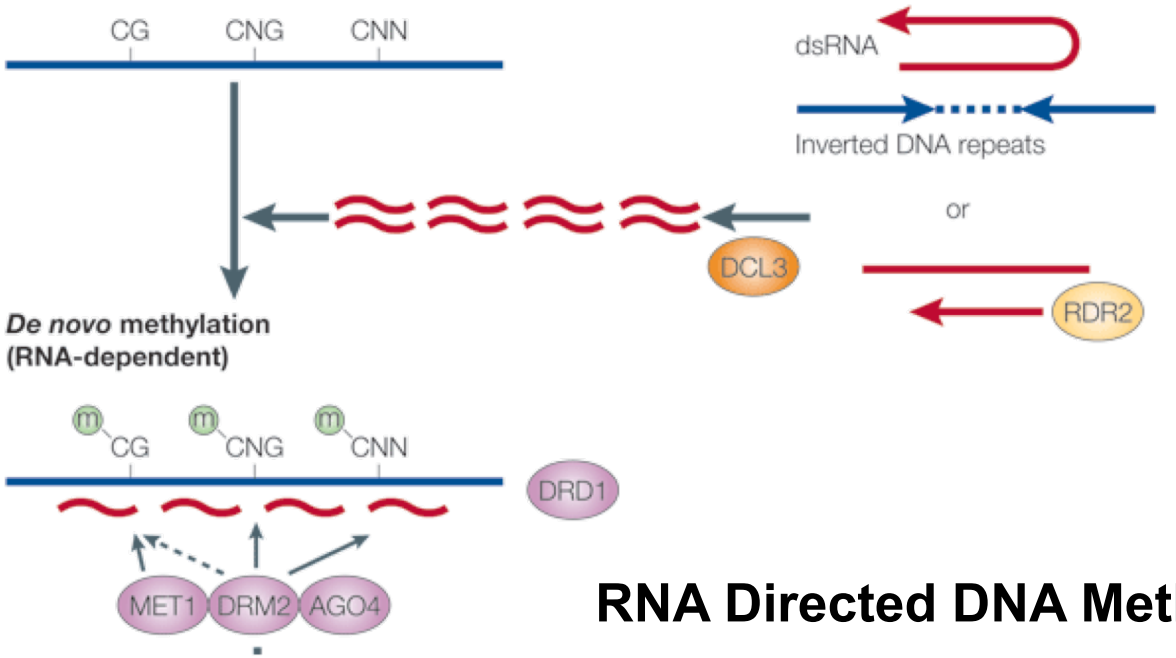
**Long ncRNA** , greater than 200nt, often much longer

**Small ncRNA**; less than 200nt, typically much shorter

microRNAs (**miRNAs**)
short interfering RNAs (**siRNAs**)
PIWI-interacting RNAs (**piRNAs**)

# Transcriptional gene silencing (TGS)

is the result of histone modifications that makes genes inaccessible to the transcriptional machinery (RNA polymerase, transcription factors, etc.).

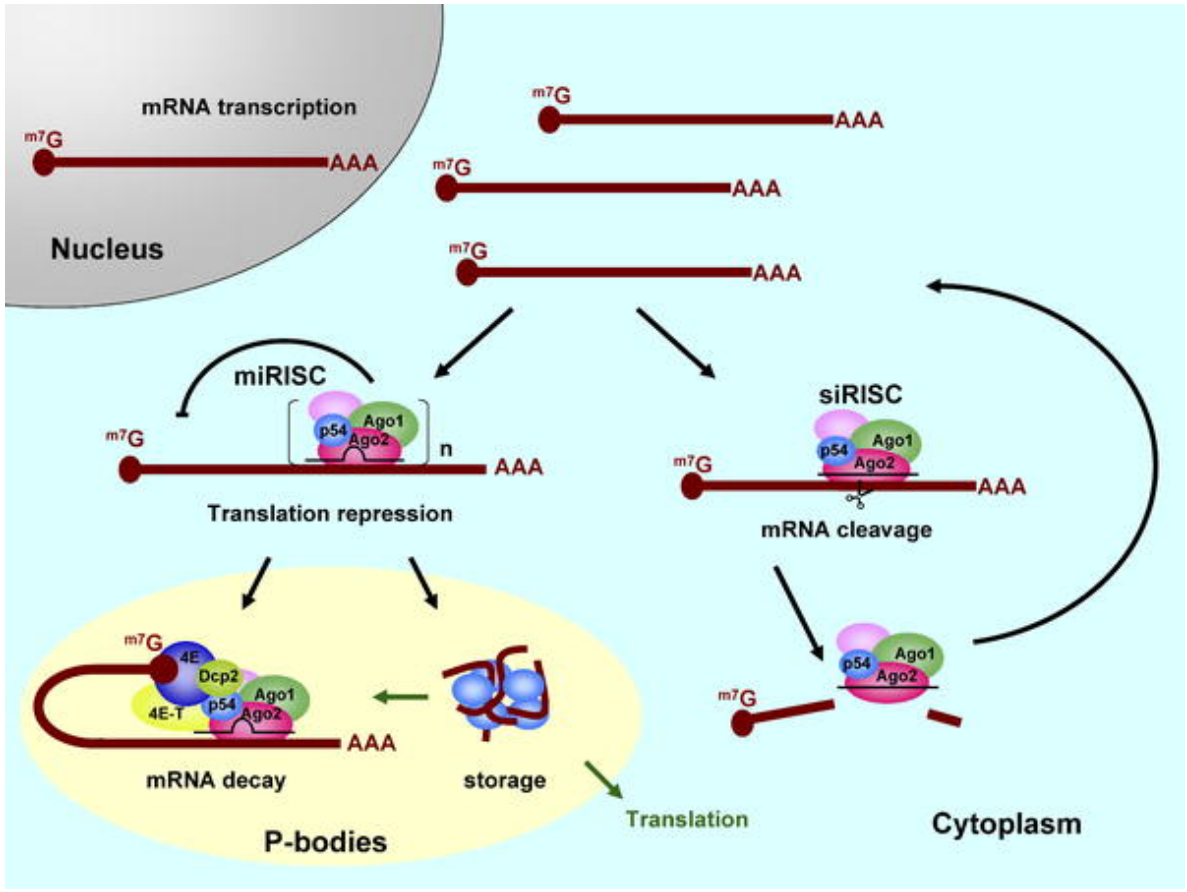



**RNA Directed DNA Methylation**

**Marjori A. Matzke & James A. Birchler**
*Nature Reviews Genetics* **6**, 24-35 (2005)

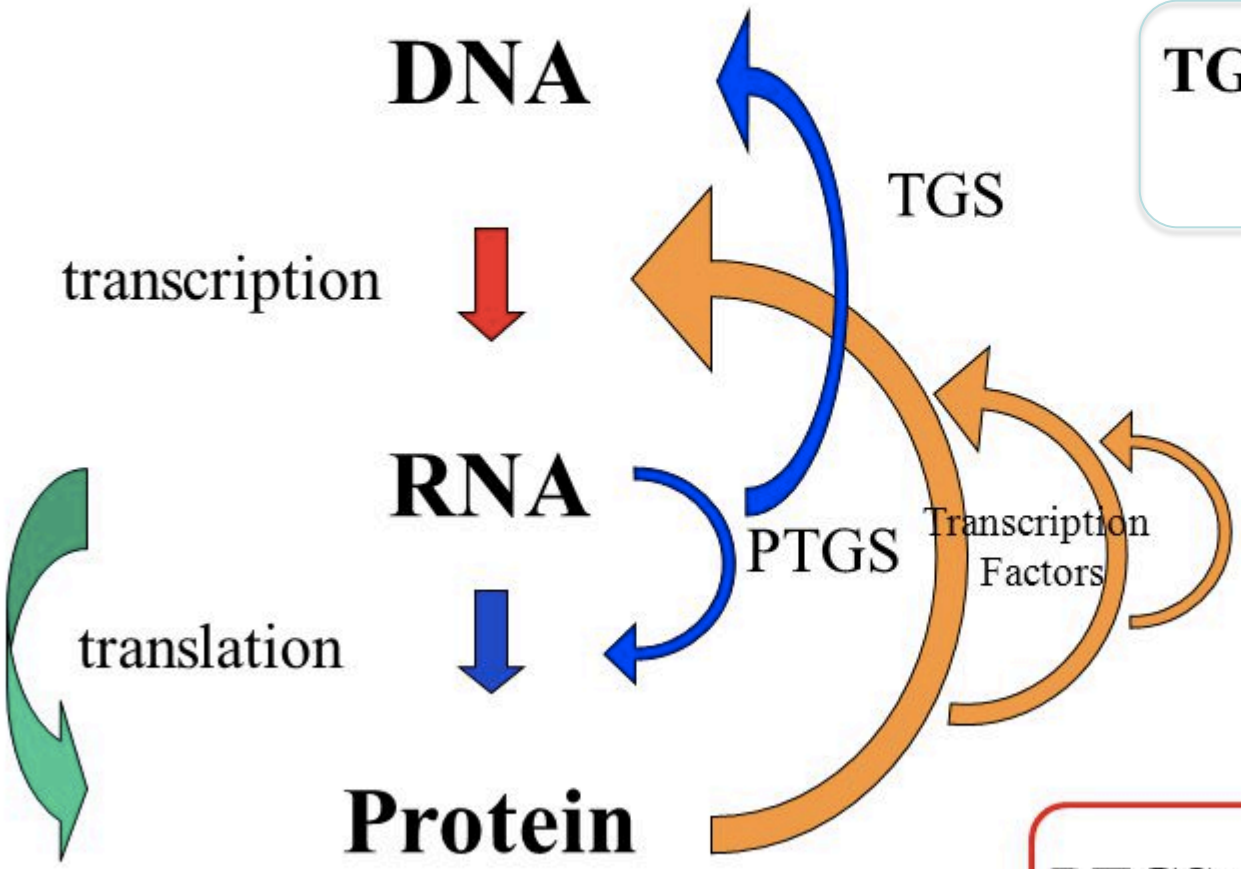
# Small ncRNA

## Posttranscriptional gene silencing (TGS)



# Central Dogma

*addenda*




**TGS:** Transcriptional Gene Silencing

**PTGS:** Post Transcriptional Gene Silencing