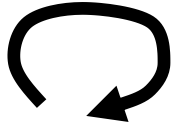
Molecular Biology 101

RNA

The Central Dogma



replication

DNA

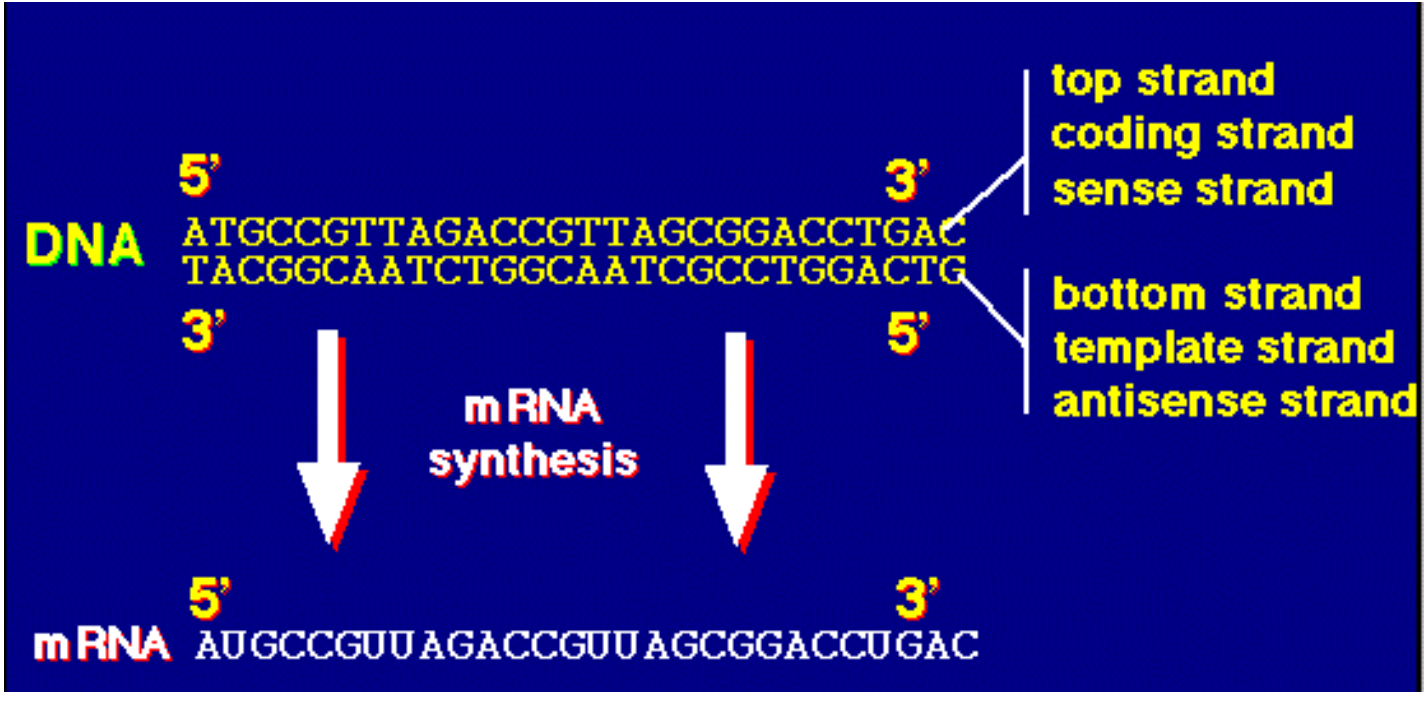
transcription

RNA

translation

Protein

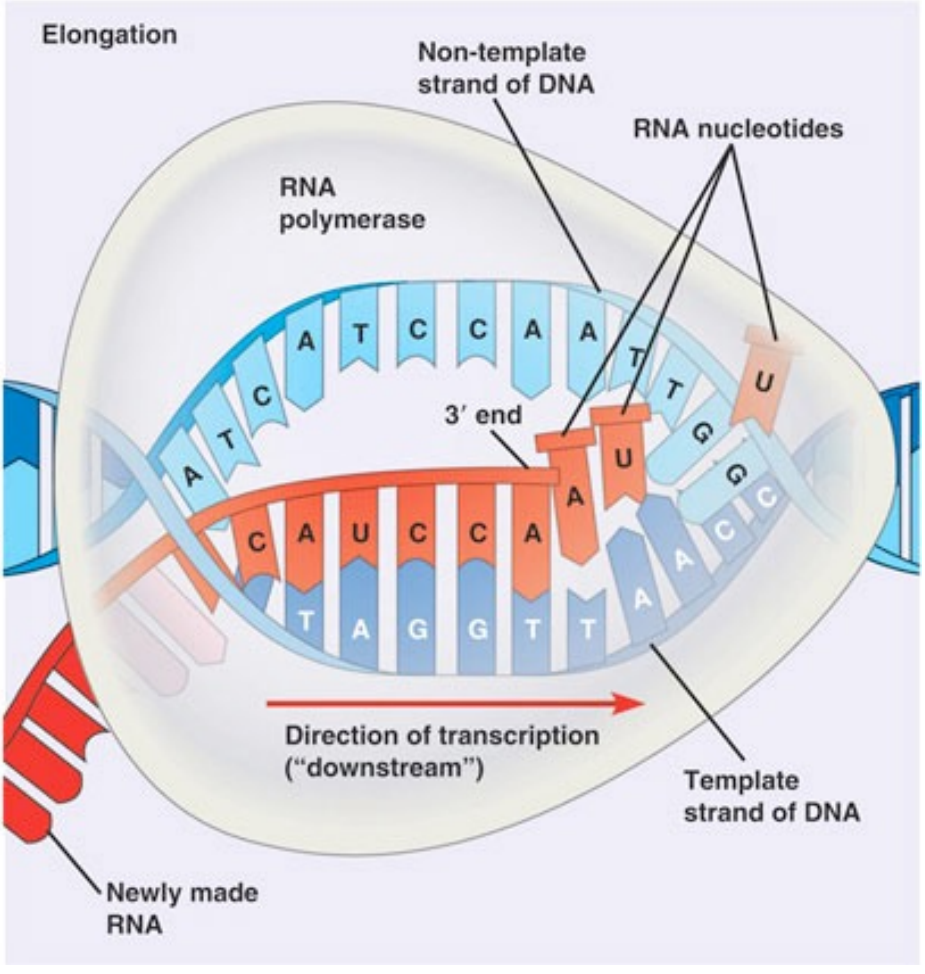
Transcription



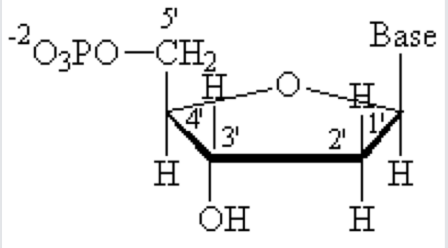
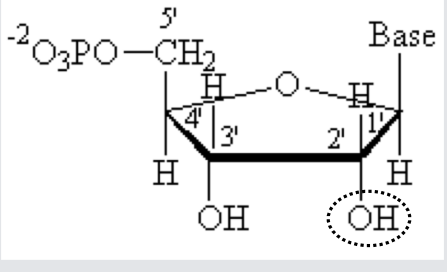
- RNA that is transcribed from a protein-coding gene is called messenger RNA (mRNA)

- RNA polymerase is the enzyme that builds an RNA molecule from a gene

Transcription: DNA→RNA



RNA vs. DNA structure

DNA	RNA
linear polymer	linear polymer
double-stranded	single-stranded
deoxyribonucleotide monomer	ribonucleotide monomer
	
A,C,G,T bases	A,C,G,U bases

The Genetic Code

Condon

Second letter

First letter	U	C	A	G	
U	UUU Phenylalanine UUC Phenylalanine	UCU Serine UCC Serine UCA Serine UCG Serine	UAU Tyrosine UAC Tyrosine	UGU Cysteine UGC Cysteine	U
	UUA Leucine UUG Leucine		UAA Stop codon UAG Stop codon	UGA Stop codon UGG Tryptophan	C A G
C	CUU Leucine CUC Leucine CUA Leucine CUG Leucine	CCU Proline CCC Proline CCA Proline CCG Proline	CAU Histidine CAC Histidine	CGU Arginine CGC Arginine CGA Arginine CGG Arginine	U C A G
			CAA Glutamine CAG Glutamine		
A	AUU Isoleucine AUC Isoleucine AUA Isoleucine	ACU Threonine ACC Threonine ACA Threonine ACG Threonine	AAU Asparagine AAC Asparagine	AGU Serine AGC Serine	U C A G
	AUG Methionine; initiation codon		AAA Lysine AAG Lysine	AGA Arginine AGG Arginine	
G	GUU Valine GUC Valine GUA Valine GUG Valine	GCU Alanine GCC Alanine GCA Alanine GCG Alanine	GAU Aspartic acid GAC Aspartic acid	GGU Glycine GGC Glycine GGA Glycine GGG Glycine	U C A G
			GAA Glutamic acid GAG Glutamic acid		

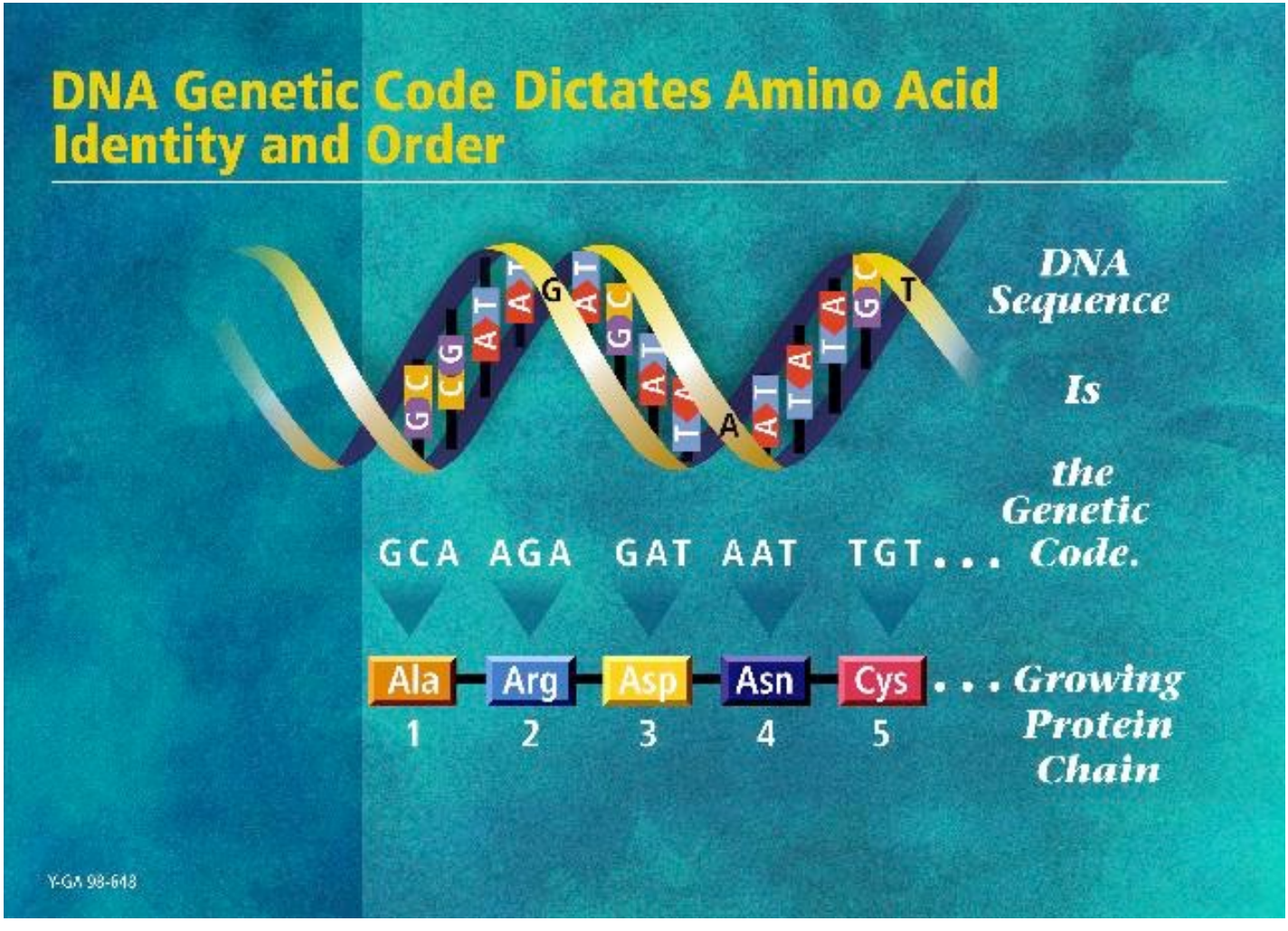
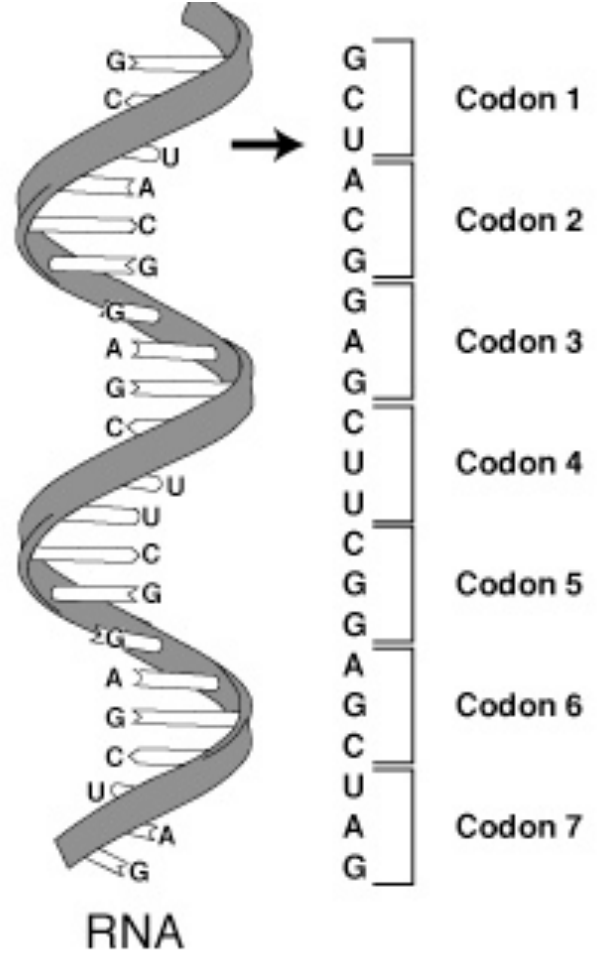


image from the DOE Human Genome Program
http://www.ornl.gov/hgmis

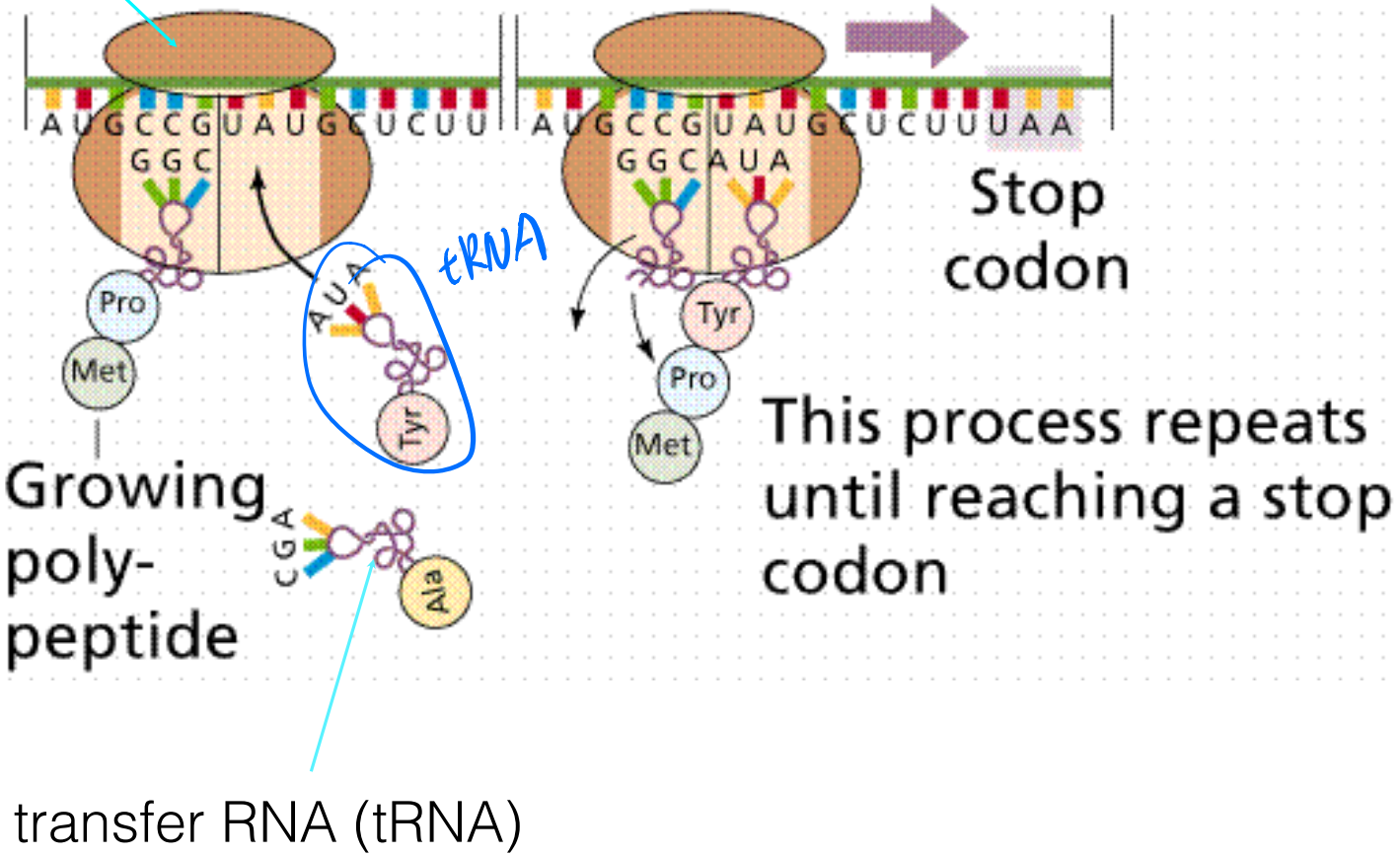
Translation

- ribosomes are the machines that synthesize proteins from mRNA
- the grouping of codons is called the reading frame
- translation begins with the start codon
- translation ends with the stop codon

Codons and Reading Frames



Translation



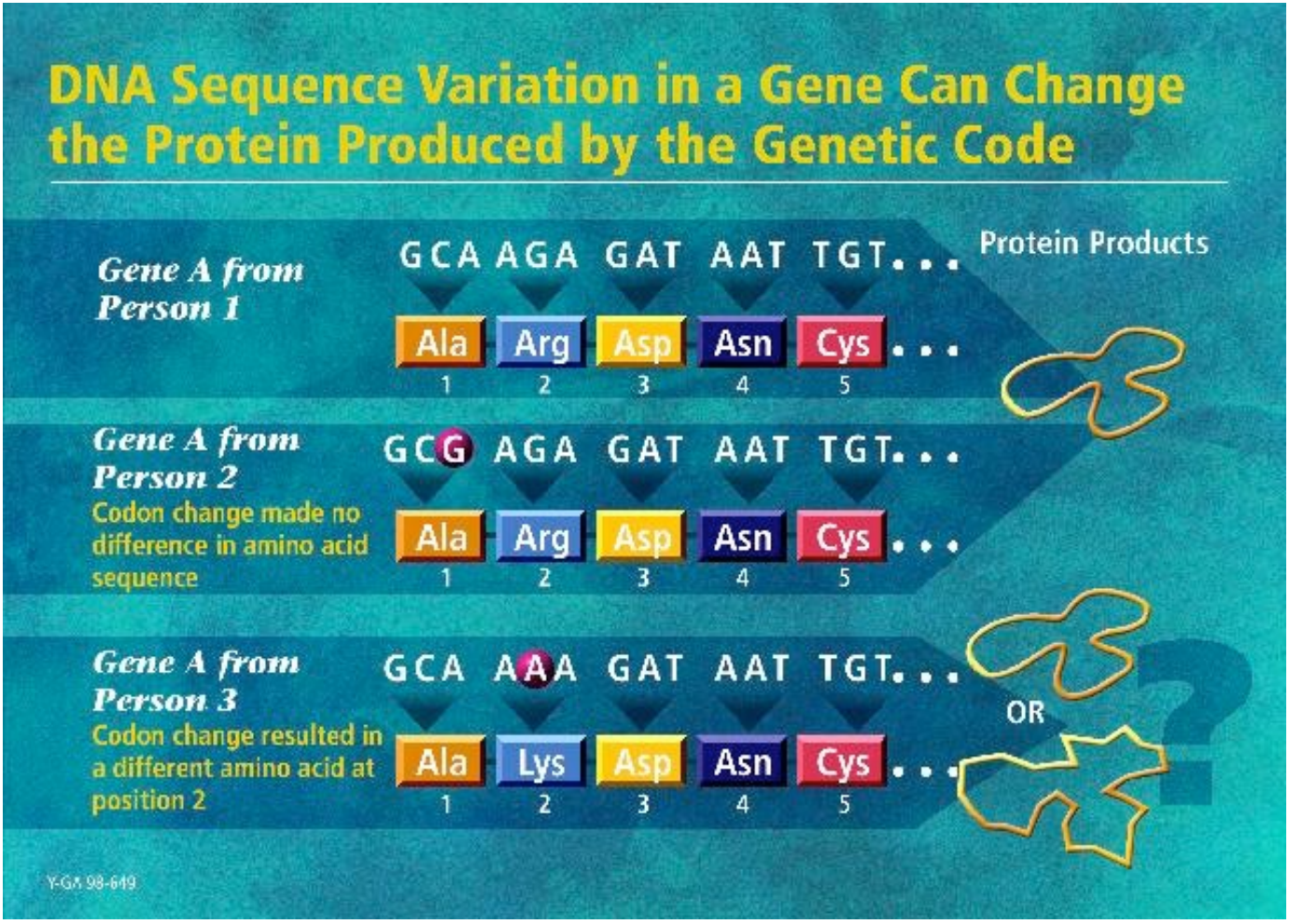


image from the DOE Human Genome Program
http://www.ornl.gov/hgmis

Genes

- genes are the basic units of heredity
- they are generally the intervals of the genome that are transcribed into RNA
- a protein-coding gene is a gene whose RNA carries the information required for constructing a particular protein (polypeptide really)
- the human genome comprises ~20,000 protein-coding genes

Gene Density

- not all of the DNA in a genome encodes protein:

bacteria	~90% coding gene/kb
human	~1.5% coding gene/35kb

RNA Genes

- not all genes encode proteins
- for some genes the end product is RNA
  - ribosomal RNA (rRNA), which includes major constituents of ribosomes
  - transfer RNAs (tRNAs), which carry amino acids to ribosomes
  - micro RNAs (miRNAs), which play an important regulatory role in various plants and animals
  - etc.