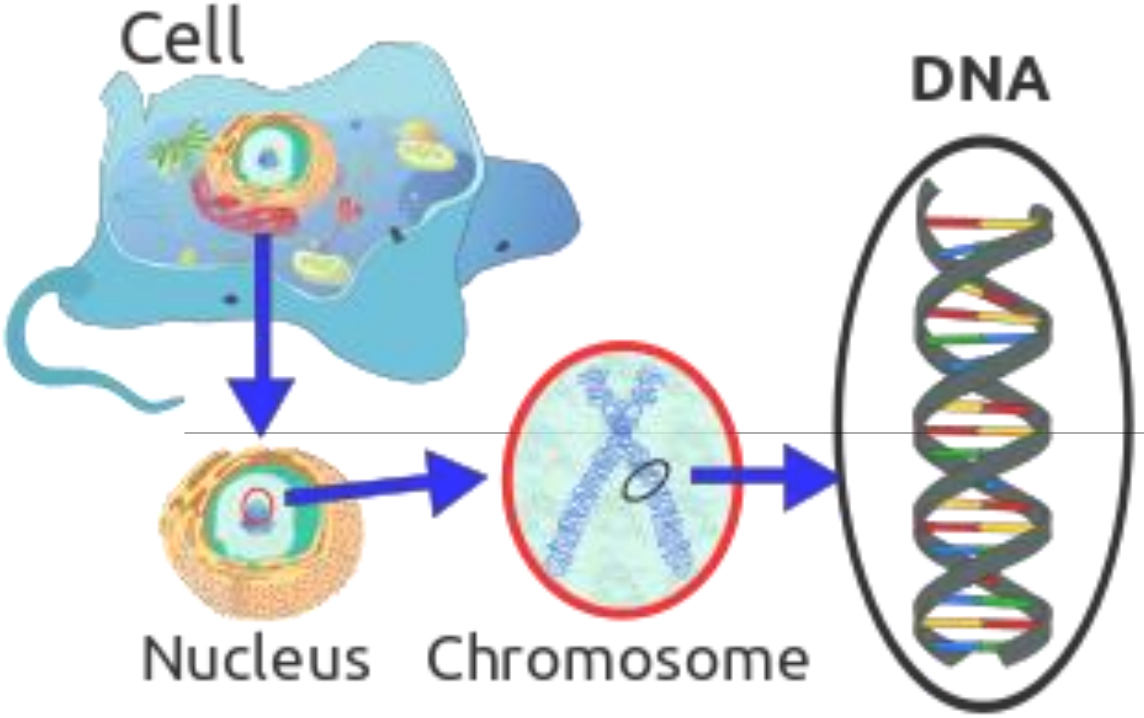
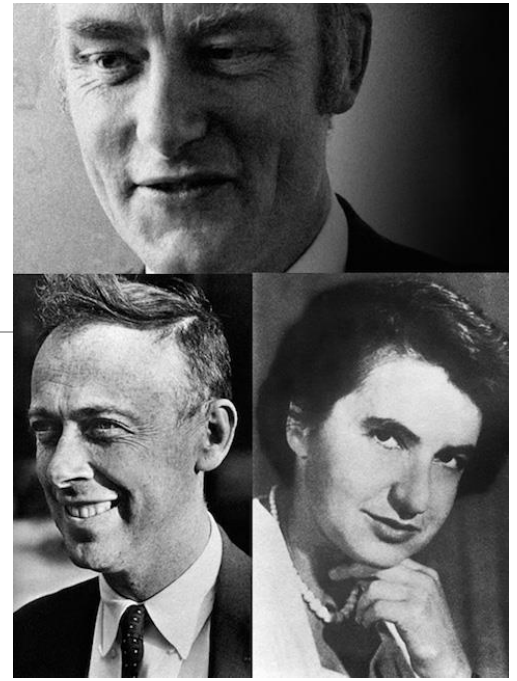
Primer on Heredity, Genes, and DNA (Chapter 1)

January 29, 2019



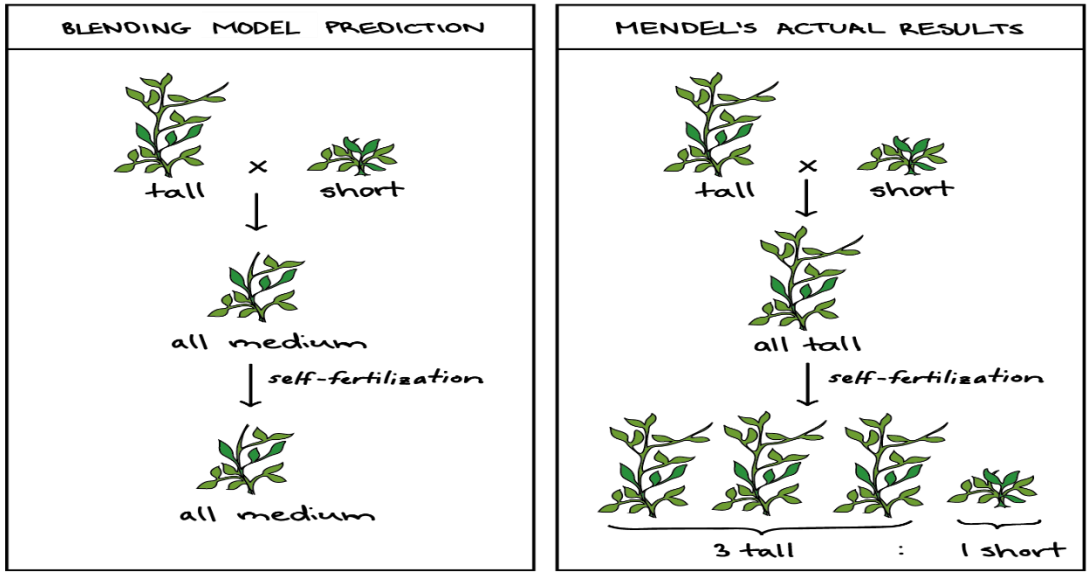
Gregor Mendel

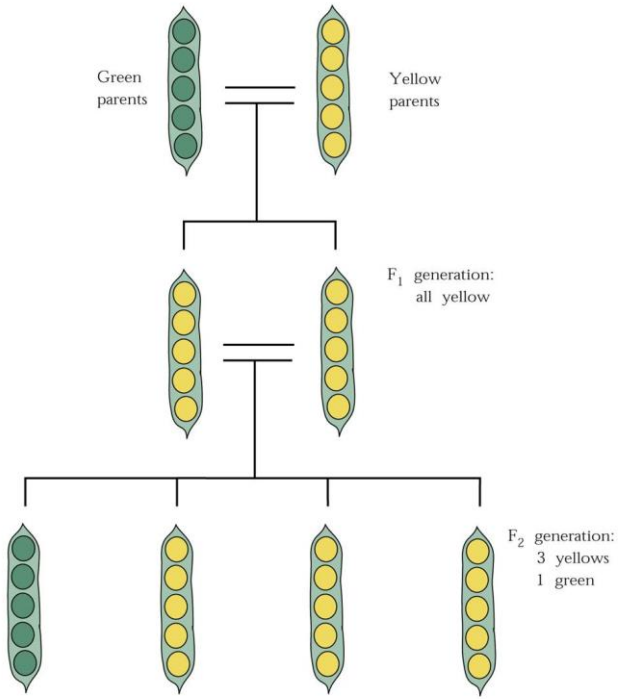
Francis Crick, James Watson, and Rosalind Franklin



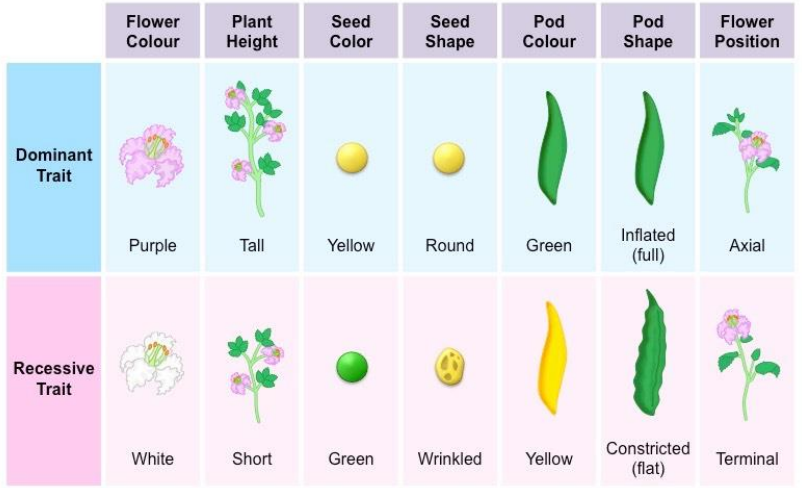
Mendelian Genetics (1850s)

- **Heredity:** transmission of information (genetic) from one generation to the next
- Mendel crossed pea plants and observed the relationship of their traits to those of their offspring
- **Key Observations**
  - **Limited** and **discrete** set of inherited traits (e.g. color and texture)
  - Traits not blended (e.g. for color: yellow or green, but NOT yellowish-green)
    - Actually always yellow (**dominant** vs. **recessive**)
  - Disappear in one generation, but could reappear in next
  - Mixing of traits from different parents

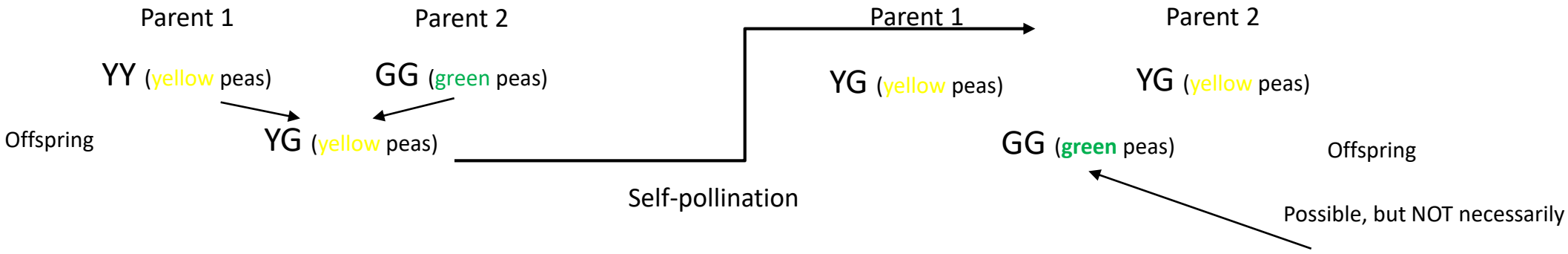




Explanation via Genes



- Traits are passed down from one generation to the next via **genes**
  - How (DNA), we'll see later
- One gene for color and another gene for texture
- Let's say 2 variants (**alleles**) each
  - Yellow (Y) and green (G) for color
  - Smooth (S) and wrinkled (W) for texture
- Note: many traits are not so discrete (continuous, or influenced by many genes, and environment)
- **Two versions of each gene** (one from each parent , chromosomes, again later)
  - If two different alleles present, **dominant** (as opposed to **recessive**) is expressed physically
  - Example, if yellow is dominant over green:



- Dominant vs. recessive is NOT absolute

What about Probabilities?

- **Independent assortment hypothesis**
  - Genes for different traits are passed down independently of one another
  - E.g.: allele for pea size does not affect allele for pea color
  - Note: NOT generally true, but may be a good approximation (later)
- Offspring gets one gene copy from father (randomly), other from mother (randomly), independent

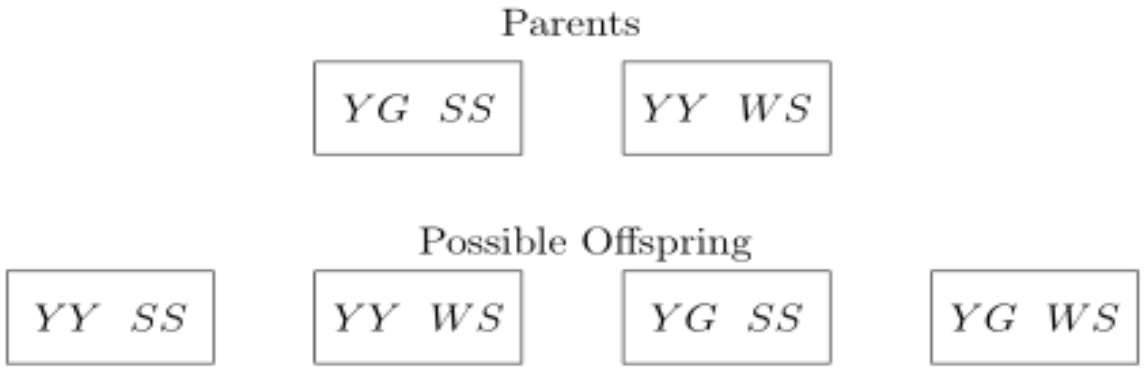


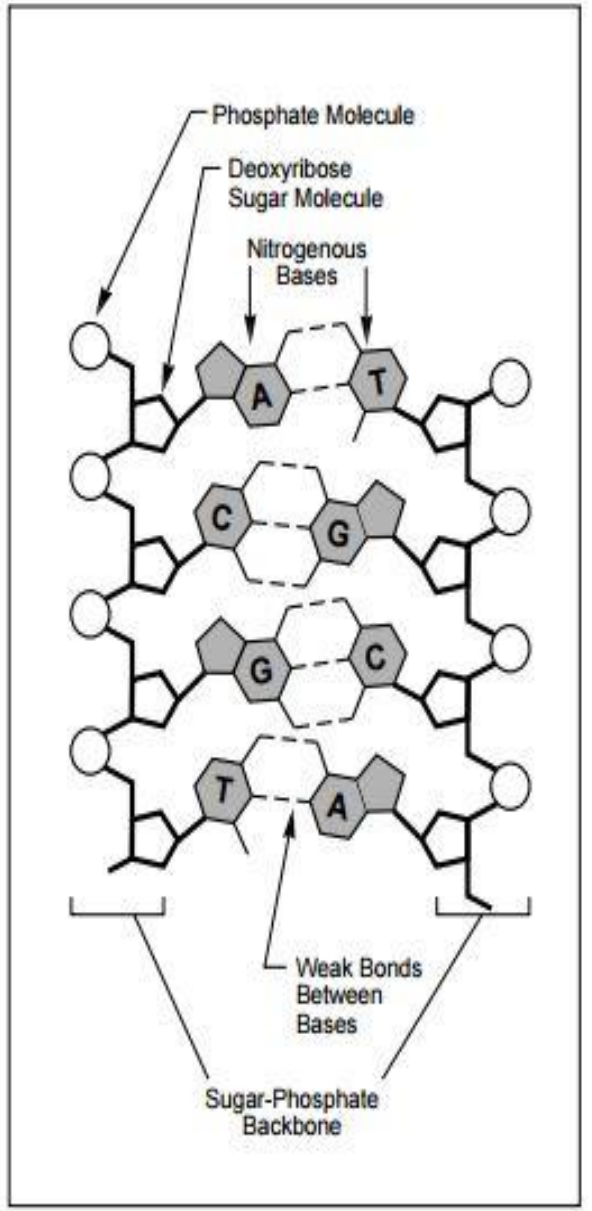
Figure 1. Assortment

- P(offspring has genes YY for color) = (1)*(1/2) = ½
- P(offspring has genes WS for texture) = (1)*(1/2) = 1/2
- P(offspring is of type YY WS) = (1/2)*(1/2) = ¼ (independence)

Genes (Biologically)

- In Eukaryotic cells, genetic material is stored in the **cell nucleus**
  - Prokaryotes – DNA processing take place in the cytoplasm itself
- Nucleus contains **genes** – basic functional unit of heredity
  - Segment of DNA used to synthesize a protein (or RNA molecule)
  - Humans – DNA organized into tightly packed structures called chromosomes (1.8 m in length)
  - 46 chromosomes in humans – 22 pairs of autosomes and a pair of sex chromosomes (XX – female, XY – male)
  - One member of each pair comes from mother (egg) and father (sperm)

Fig. 2. DNA Structure. *The four nitrogenous bases of DNA are arranged along the sugar-phosphate backbone in a particular order (the DNA sequence), encoding all genetic instructions for an organism. Adenine (A) pairs with thymine (T), while cytosine (C) pairs with guanine (G). The two DNA strands are held together by weak bonds between the bases. A gene is a segment of a DNA molecule (ranging from fewer than 1 thousand bases to several million), located in a particular position on a specific chromosome, whose base sequence contains the information necessary for protein synthesis.*



Chromosomes and karyotype

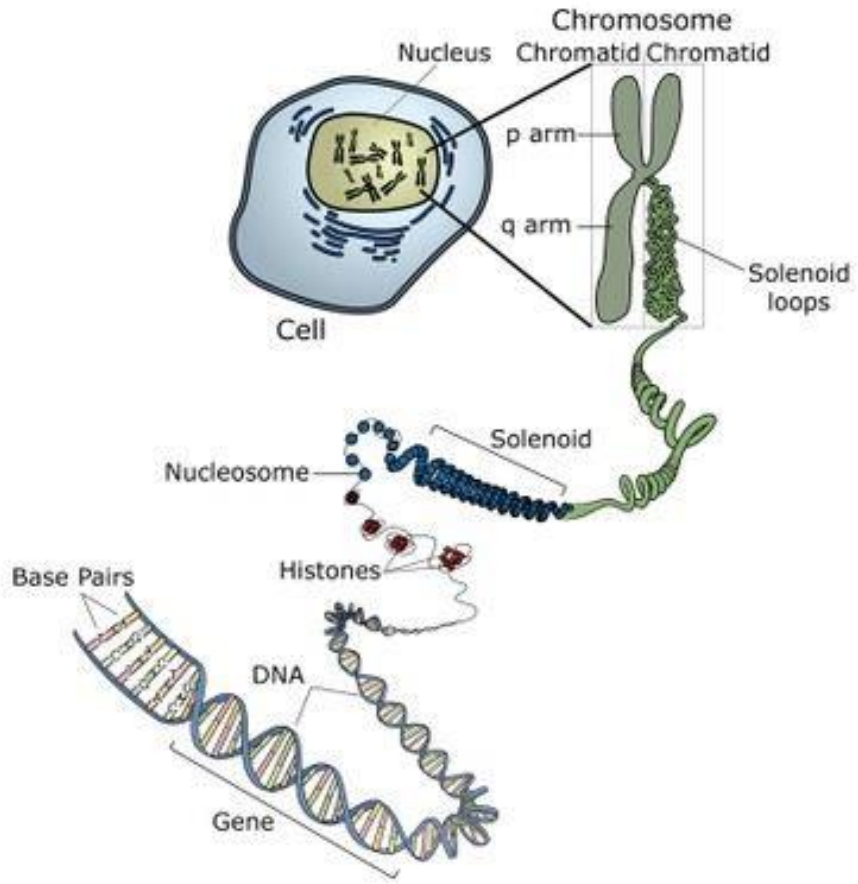
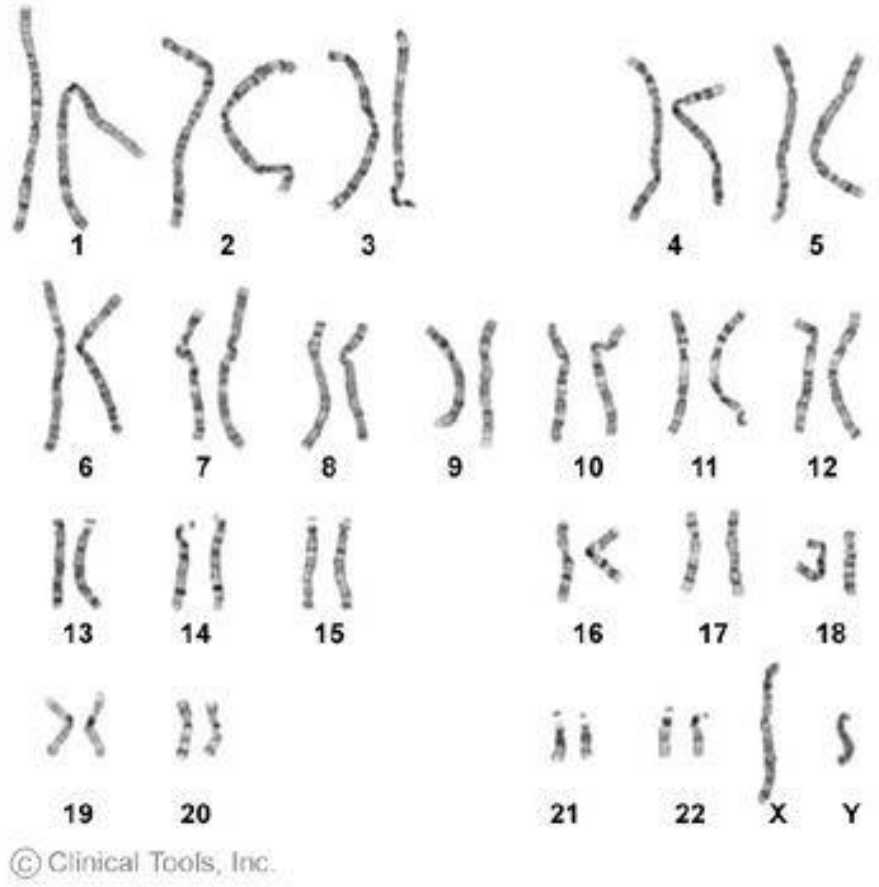


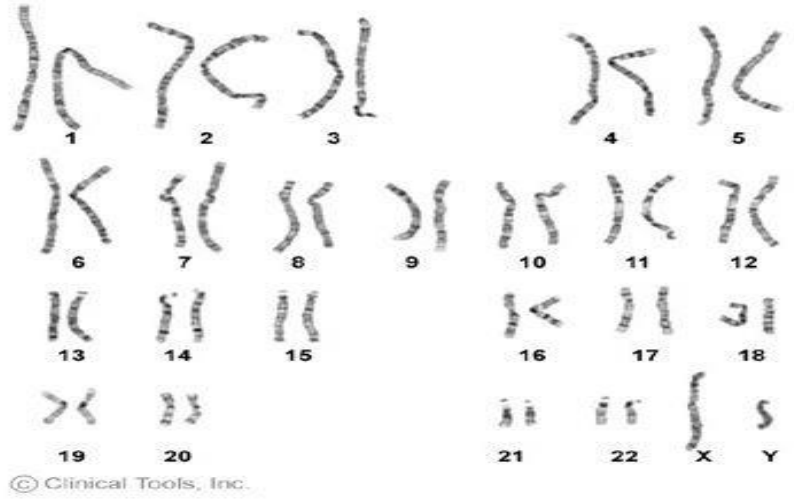
Image adapted from: National Human Genome Research Institute.



Karyotype of a human male

Physical Mechanism of Mendel's Heredity in DNA

- **Chromosomes** are strands of DNA, and **genes** are small subintervals of a strand
  - Portion of chromosome where gene is located is called its **locus**
  - **Alleles:** alternate forms of gene that can reside at a locus
- **Diploid:** chromosomes come in (possibly distinct!) pairs, except maybe sex chromosome (e.g. Mendel's peas, YY or YG or GG)
- **Haploid:** one copy of each chromosome
  - Note: males only have one X chromosome, while females have two
  - **Polyploidal:** 2n copies, n > 1

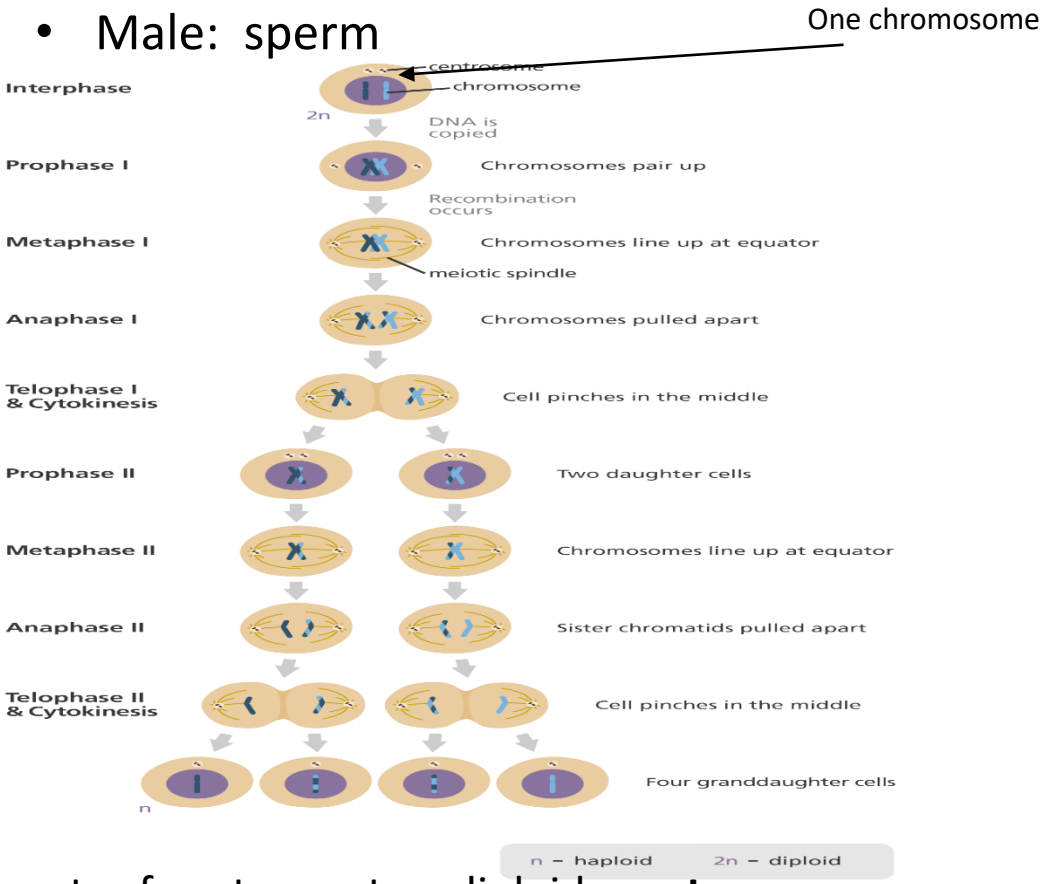


Karyotype of a human male

Karyotype of human female

Passing of Genetic Information (Meiosis)

- Diploid organisms produce haploid cells (**gametes**) for sexual reproduction
  - Female: eggs
  - Male: sperm



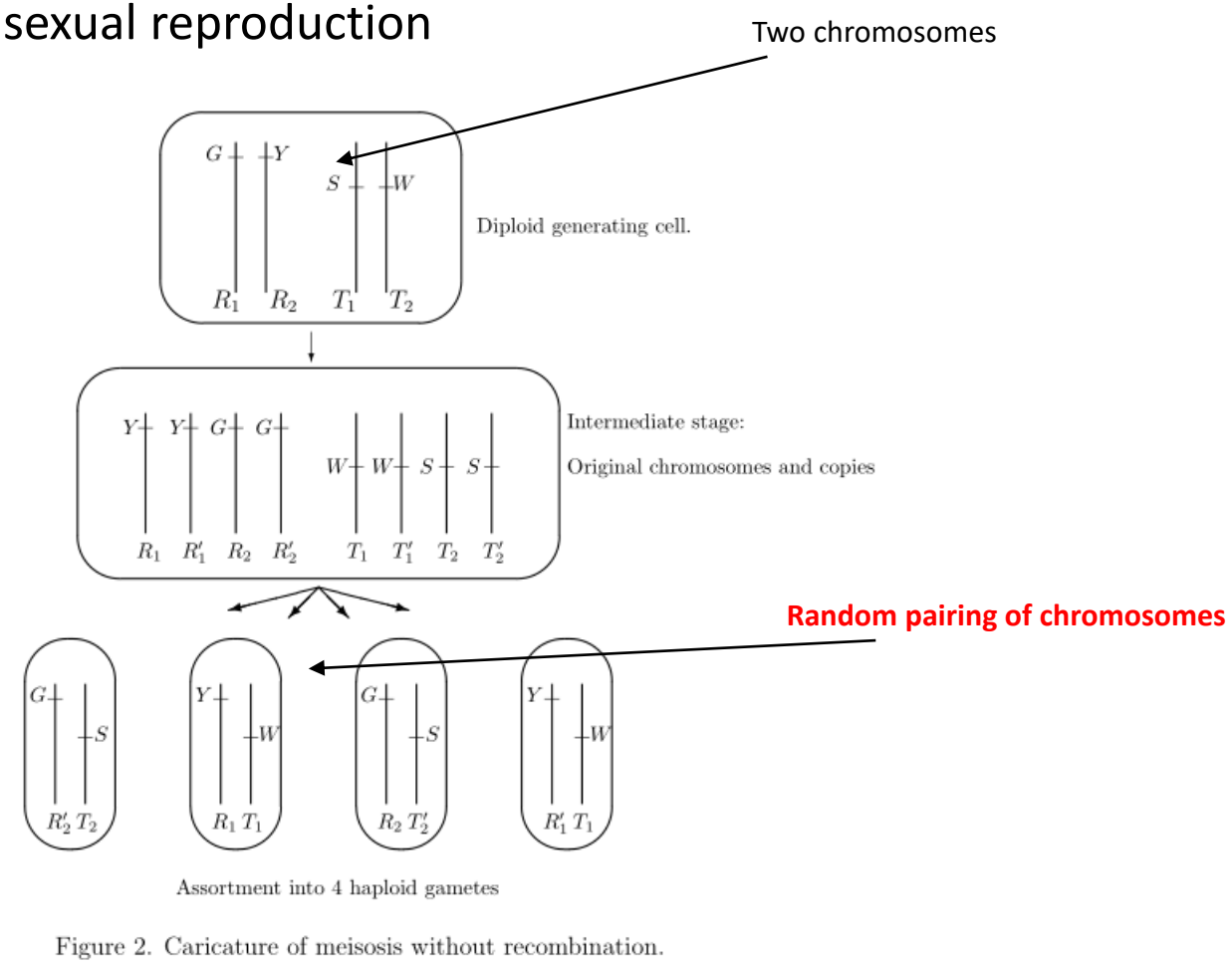


Figure 2. Caricature of meisosis without recombination.

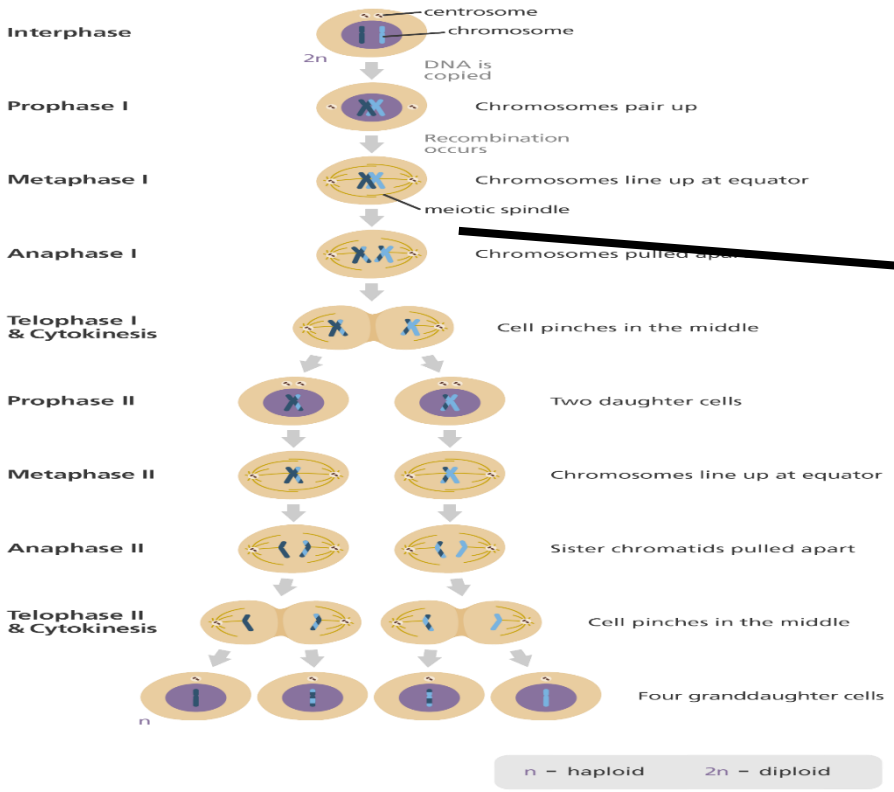
- Gametes fuse to create a diploid **zygote**
  - **Homozygote:** two identical alleles of a specific gene
  - **Heterozygote:** different alleles at a gene

See why independent assortment may fail if genes on same chromosome

Genotype: list of alleles it carries at the loci of its chromosome

Recombination

- Alleles on same chromosome may not be strictly linked because of **recombination**
  - Before being separated, duplicate chromosomes form a complex and may exchange genetic material
  - Novel set of genetic information that can be passed on from the parents to the offspring.



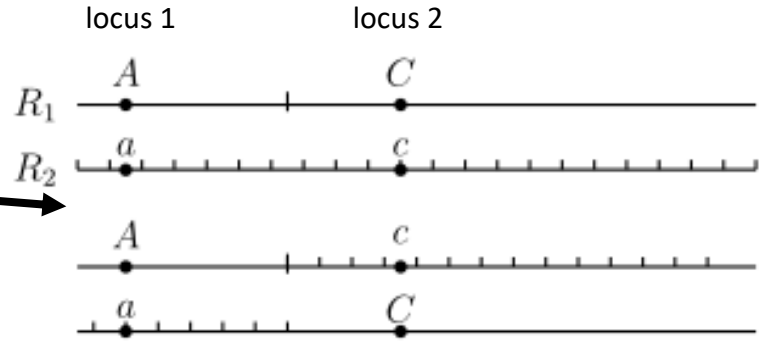
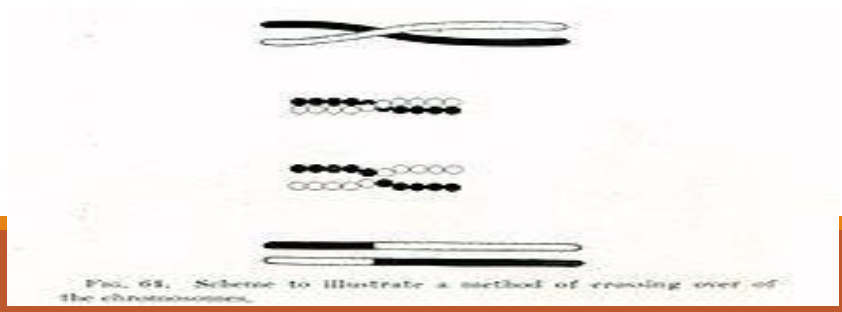


Figure 3. Interchange of chromosomal material in recombination

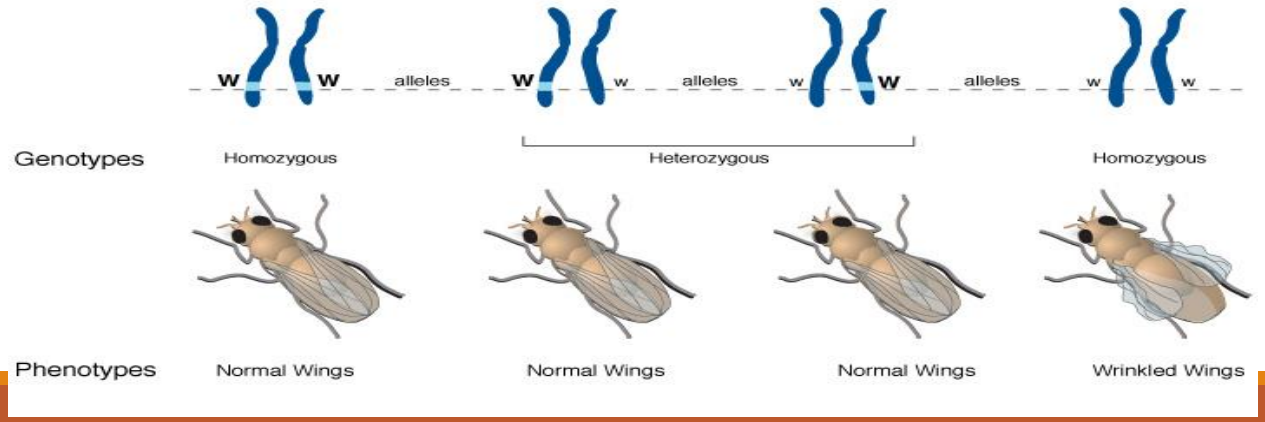
Recombination frequency: degree to which alleles are linked

- Fundamentally linked to how close genes (loci) along a chromosome

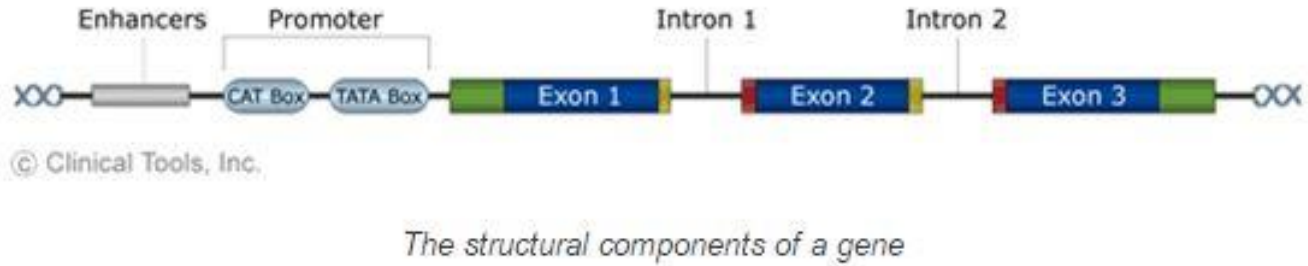


Genotype and Phenotype

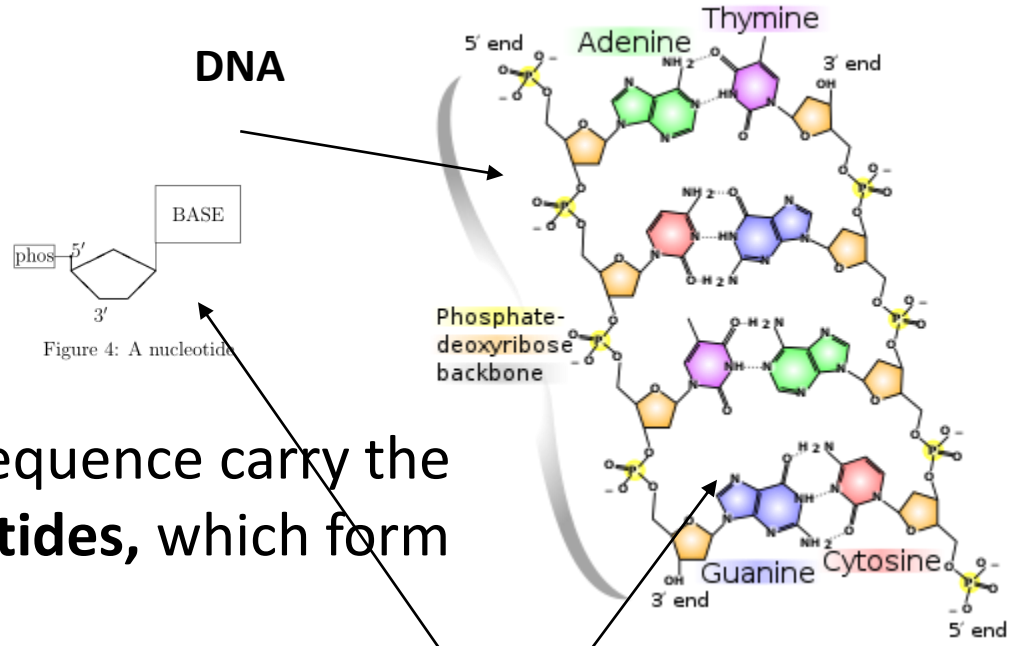
- **Genotype:** list of alleles it carries at the loci of its chromosome
  - With respect to a (small) number of traits
  - **Color:** YY, YG, and GG
  - **Color AND texture:** 9 options (YY SS, YY SW, YY WW, YG SS, YG SW, YG WW, GG SS, GG SW, GG WW)
- **Phenotype:** set of organisms physical, biochemical, and behavioral traits (how it looks and functions)
  - Relationship between genotype and phenotype is complicated and **NOT** understood (yet!)
  - **Different genotypes may present same phenotype**
    - Dominance of alleles
    - Environment
    - Interaction of many genes on a single trait
    - **Epigenetics:** DNA methylation and histone modification may affect gene expression



Genes and Proteins



The structural components of a gene



- Genes – specific sequence of nucleotide bases, whose sequence carry the information required for constructing **proteins (polypeptides,** which form proteins)
  - Proteins – provide structural components of cells and tissues, as well as enzymes
    - Large, complex molecules made up of amino acids
    - Synthesize over 100,00 different kinds
  - Humans – over 100,000 genes and 3 billion base pairs in 23 chromosomes
- In a gene, codons direct for a specific sequence of amino acids to produce proteins
  - *Expression of gene* means production of the protein it codes for

Each element is a **nucleotide**

Think if DNA a strings of letters from DNA alphabet {A,G,C,T}

These strings code for **proteins**

Genes and Proteins continued

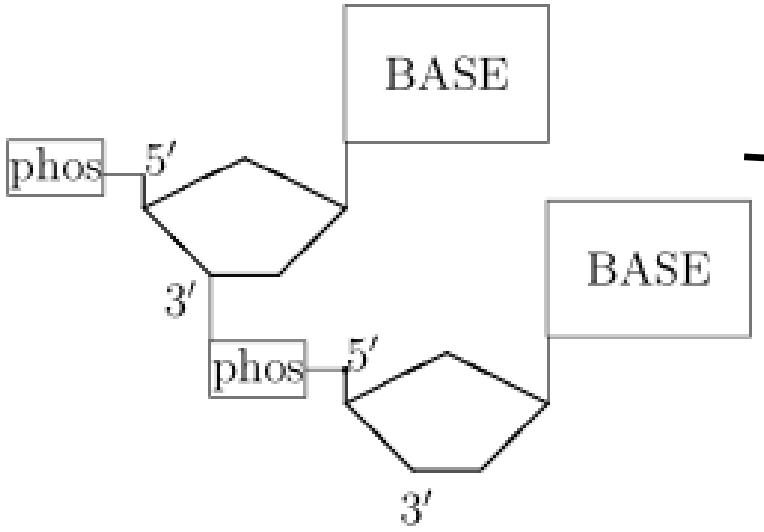


Figure 5: Linked nucleotides

$5' - TAGGTTAGGCTATTAGGCTGA - 3'$

Double helix structure

Base pair (*bp)*

$5' - TAGGTTAGGCTATTAGGCTGA - 3'$

$3' - ATCCAATCCGATAATCCGACT - 5'$

Transcription and Translation

- Process by which genes, **when expressed**, are used in the synthesis of a functional gene product (usually proteins)
- "Central Dogma of Molecular Biology"

Every 3 letters of DNA (**codon**) determines 1 amino acid

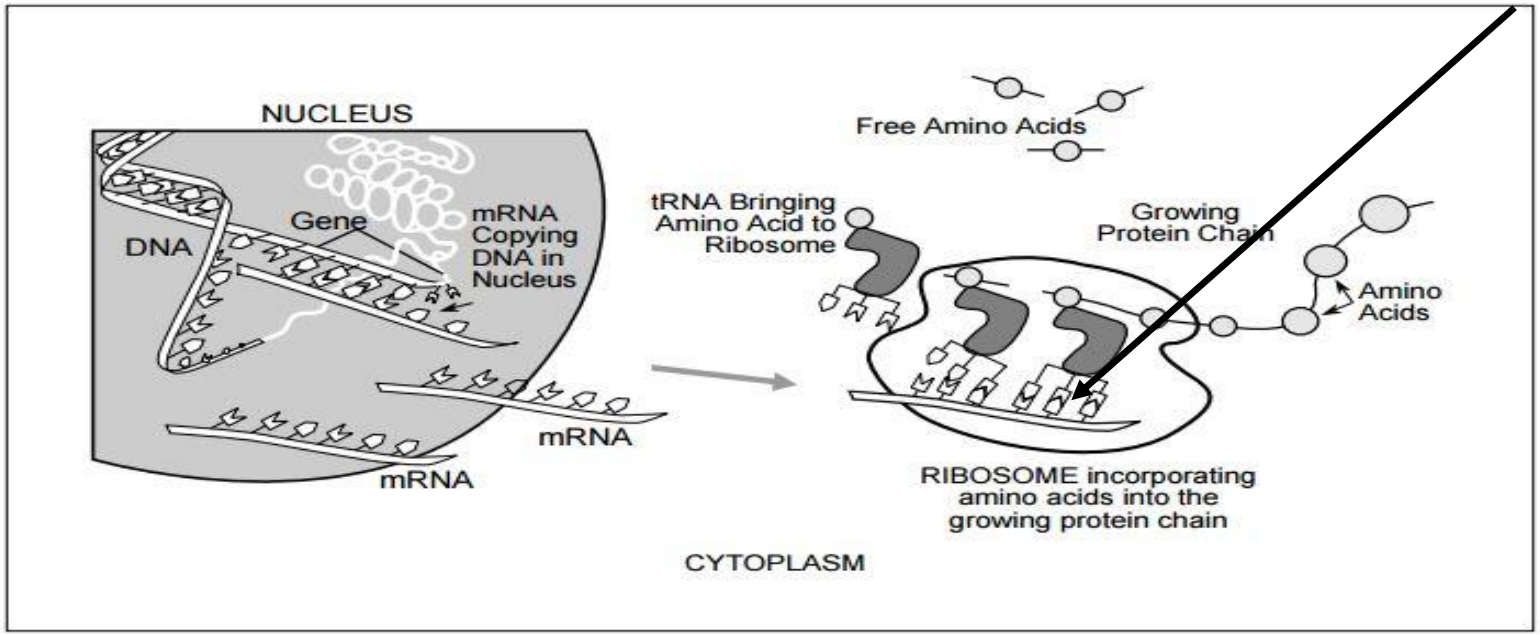
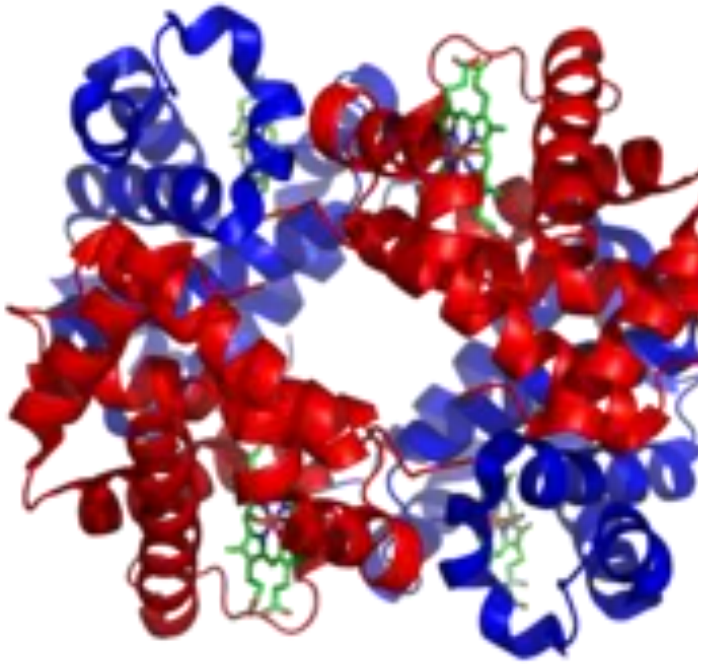


Fig. 5. Gene Expression. *When genes are expressed, the genetic information (base sequence) on DNA is first transcribed (copied) to a molecule of messenger RNA in a process similar to DNA replication. The mRNA molecules then leave the cell nucleus and enter the cytoplasm, where triplets of bases (codons) forming the genetic code specify the particular amino acids that make up an individual protein. This process, called translation, is accomplished by ribosomes (cellular components composed of proteins and another class of RNA) that read the genetic code from the mRNA, and transfer RNAs (tRNAs) that transport amino acids to the ribosomes for attachment to the growing protein. (Source: see Fig. 4.)*

Amino Acids and Proteins

- Proteins are long strings of amino acids which fold into a specific shape
  - Code and shape determines the precise structure (folded 3-dimensional structure is key to understanding function)



Amino acid	Abbr.	Codons
Alanine	A	GC*
Arginine	R	AGA,AGG,CG*
Asparagine	N	AAT, AAC
Aspartic acid	D	GAT, GAC
Cysteine	C	TGT, TGC
Glutamic acid	E	GAA, GAG
Glutamine	Q	CAA, CAG
Glycine	G	GG*
Histidine	H	CAT, CAC
Isoleucine	I	ATT, ATC, ATT
Leucine	L	CT*, TTA, TTG
Lysine	K	AAA, AAG
Methionine	M	ATG
Phenylalanine	F	TTT, TTC
Proline	P	CC*
Serine	S	TC*, AGT, AGC
Threonine	T	AC*
Tryptophan	W	TGG
Tyrosine	Y	TAT, TAC
Valine	V	GT*

Table 1.1: Amino Acids and the Genetic Code

*: any base (A,C,T,G) may occupy

- **Start:** TAA, TAG
- **Stop:** TGA