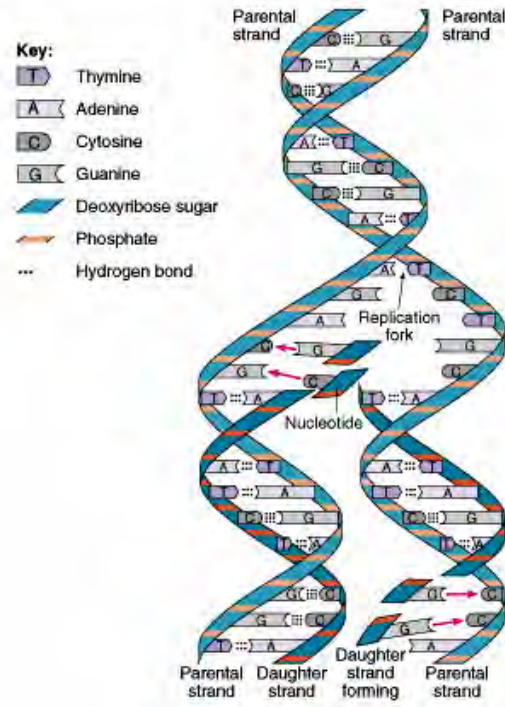


# MOLECULAR EVOLUTION

## Deoxyribonucleic acid (DNA)



Double helix, each strand made up of nucleotide sequences

Nucleotides are composed of sugar, phosphate, and base

There are four different bases in DNA:

Adenine (A)

Thymine (T)

Cytosine (C)

Guanine (G)

Each base will pair with only one of the others:

A – T, C – G

In other words, the pattern in which bases pair is specific.

Base-pairing specificity:

1. Determines structure of the double helix
2. Allows DNA to replicate itself

The genetic code is based on the sequence of nucleotides on a DNA strand, which ultimately leads to the structural and functional components of cells; i.e., proteins:

1. The genetic code uses the sequence of nucleotides in DNA to code for proteins.
2. Proteins are complex, large, chain-like molecules that act as enzymes (catalysts) and structural components in cells.
3. All proteins are composed of long sequences of (20 naturally occurring) amino acids linked together in the chain.
4. There are many different proteins because each sequence of amino acids is a different protein. Change the sequence; change the protein.

Back to DNA and the genetic code:

1. A codon is three consecutive DNA nucleotides, each of which (more or less) codes for an amino acid.
2. How many different codons are there? In other words, if there are four different nucleotides (A, T, C, G), how many different combinations of three of them are possible? The answer is  $4 \times 4 \times 4 = 64$  possible ways that the 4 different kinds of nucleotides can be grouped in sequences of 3.

- The order of codons (or more fundamentally, the order of nucleotides arranged in groups of three) in a DNA strand determines (in other words, codes for) the order amino acids that comprise a specific protein.
- Because there are only 20 naturally occurring amino acids, yet there are 64 different codons, and except for some special codons, each codon codes for an amino acid, most amino acids can be specified by several different codons.
- In that sense, the genetic code is redundant.

TABLE 1. THE DNA CODONS AS PRESENTLY KNOWN.  
ATT, ATC, AND ACT DO NOT CODE FOR ANY AMINO ACID

AAA } AAG }	Phenylalanine	AGA } AGG } ACT } ACC }	Serine	ATA } ATG }	Tyrosine	ACA } ACC }	Cysteine
AAT } AAC }	Leucine			ATT } ATC }	Stop	ACT	Stop
CAA } GAG } CAT } CAC }	Leucine	GCA } GCC } CGT } CCC }	Proline	GTA } GTC }	Histidine	ACC	Tryptophan
TAA } TAG } TAT }	Isoleucine	TGA } TGG } TGT } TGC }	Threonine	GTT } GTC }	Glutamine	<del>GCA</del> <del>GCC</del> CCT } CCC }	Arginine
TAC	Methionine			TTA } TTG }	Asparagine	TCA } TCG }	Serine
CAA } CAG } CAT } CAC }	Valine	CGA } CCC } CGT } CCG }	Alanine	TTT } TTC }	Lysine	<del>TCT</del> TCC }	Arginine
				CTA } CTC }	Aspartic acid	CCA } CCG } CCT } CCC }	Glycine
				CTC } CTT }	Glutamic acid		

Note that most often, in cases where the same amino acid is coded for by several codons, the difference in the codons is in the third nucleotide of the sequence.

### Mutations

Definition: Changes in nucleotide sequence; hence, changes in genes  
Mutations may have no effect, slight or harmless effects, or may be disastrous

## Kinds of mutations

1. Chromosome mutation  
Addition, deletion, or scrambling the order of nucleotides;  
Frame shift
2. Chromosome mutations usually have profound effects while point mutations do not necessarily have profound effects.
3. Point mutation  
Substitution of a single DNA nucleotide by another

### Point mutations:

#### Silent point mutations:

1. A point mutation may be silent in the sense that the nucleotide substitution does not change the amino acid coded for;
2. Most silent mutations involve substitution at the third nucleotide position;
3. Almost all substitutions at the third nucleotide position are silent.

#### Replacement point mutations:

1. A nucleotide substitution changes the amino acid coded for;
2. A replacement point mutation may not change the structure or function (or both) of the protein of which it is a part; or
3. A replacement point mutation may have structural or functional effects (or both), usually bad.

Neutral mutations are silent or occur in non-functional, non-critical portions of proteins; they do not affect the fitness of the individual.

Mutations add variability to the genome.

Mutations occur as a cell prepares to divide. DNA is copied so that each daughter cell will contain a complete set of genetic instructions. Proof-reading enzymes compare the copy with the original and correct

copying errors in most cases, but occasionally the copy is not faithful to the original, resulting in a mutation.

Thus, the rate of mutation is in a sense controlled by specific enzymes, which are themselves coded for by specific genes.

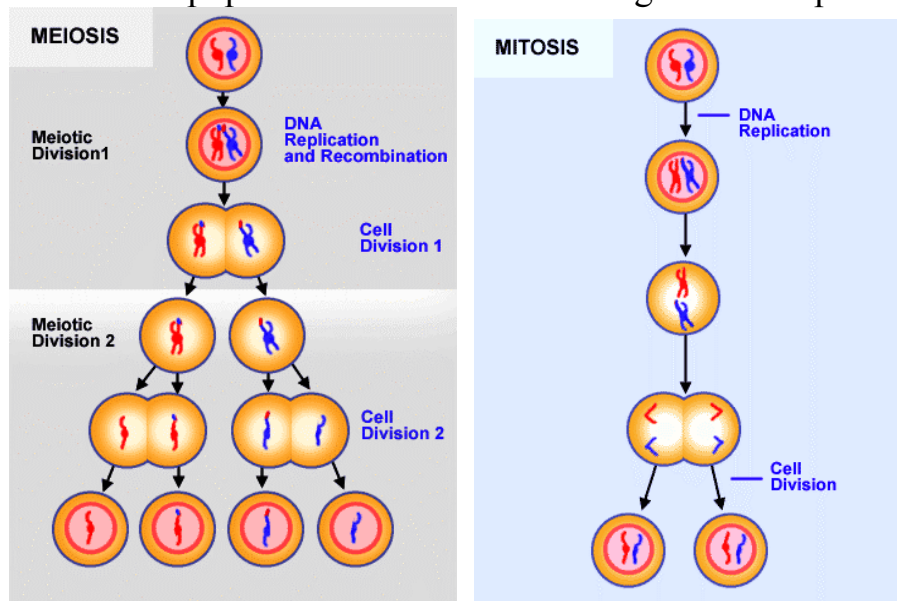
Moreover, even when DNA is not being duplicated, it is being scanned for damage sustained in normal cellular wear and tear, and the damaged parts may be repaired.

Experiments with bacteria have shown that when environmental conditions are unfavorable, the tendency to mutate is great; when conditions are more favorable, the tendency to mutate is greatly reduced.

Mutations may occur at any position of many positions on the chromosomes; i.e., in many different genes. Which particular gene will mutate appears to be random. Any gene not affected by the mutation-suppressing enzyme has an equal probability to mutate.

Unsuppressed, random mutation then adds (potentially) to individual variation on which natural selection may act, consistent with Darwinian evolution.

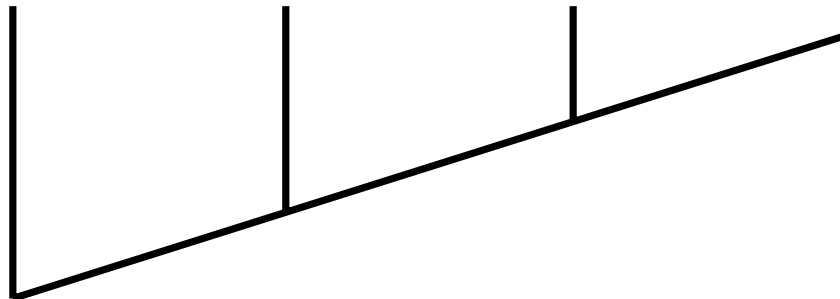
Genetic variation in populations is enhanced through sexual reproduction:



## Evolutionary information from molecules:

Long chain molecules (such as DNA and proteins) made up of a succession of discrete moieties (nucleotides in DNA; amino acids in proteins) contains evolutionary information simply because of their structure and the way their structure is modified through mutations. Mutations happen in time, and they can only happen to existing molecules. Once a mutation occurs, the structural change in a DNA molecule is maintained. Over time, the changes add up. The cumulative amount of change increases with time. The degree of difference in DNA sequences can be compared among evolutionary lineages, that is, among living species, to gain an understanding of which species are more similar in their DNA structure, and hence, which are more closely related genetically and evolutionarily. The more similar the DNA, the more closely related the species.

Degree of relatedness is reflected in the numbers of nucleotide sequence differences among DNA of different species. The pattern for some primates is:



This pattern implies (even without resorting to the fossil record) that humans and gorillas shared a common ancestor after humans and orangutans shared an ancestor, and all three shared a common ancestor not shared with gibbons.

This figure shows a phylogeny, or pattern of relationships based on relative genetic closeness as determined from the sequence of inferred common ancestry. Genetic closeness is determined by the similarity of DNA nucleotide sequences. The pattern shows who is more closely related to whom and it implies a time sequence.