Chapter 8

THE ORIGIN OF GENETIC VARIATION

IMPORTANT: Review chapters 15 to 18 in your Principles of Biology I textbook, Campbell 9th Ed.

Also: http://www.dnaftb.org/

The fundamental process of evolution is a change in the inherited characteristics of a population or species.

Genetic changes in populations and species begin with changes in the genetic material carried by individual organisms: mutations.

GENES AND GENOMES

The genome is total amount of DNA in the cell of an organism. The entire set of chromosomes of an organism.

- The genome contains coding and non-coding DNA.

The genome contains all the instructions necessary to make an organism and its growth and development, all of which is controlled by the genes within the genome.

A distinction may be made between the nuclear genome and organelle genomes.

In the case of some viruses, the genome is made of RNA.

DNA is made up of base pairs nucleotides, bp, one adenine or guanine, and one thymine or cytosine: A-T and C-G.

The DNA content varies greatly among organisms:

- The haploid genome of Drosophila melanogaster is about 1.5 x 10⁸ base pairs (bp).
- That of humans is about 3.2 x 10⁹ bp, 3.2 billion base pairs.
- The genome of Amoeba dubia is about 200 times the size of the human genome.
- Humans have about 20,000 genes although some estimates go as high as 100,000.

The genome of an organism contains genes and non-coding DNA.

A gene is a portion of DNA that is transcribed into RNA.

- Gene is functional definition. It is the hereditary unit found in a chromosome.
- A gene contains all the instructions to make a specific protein molecule.
- Locus is the site in the chromosome occupied by a particular gene.
- Thousands of genes in the human genome encode for tRNA and rRNA that are never translated into proteins.

One strand of a protein-encoding gene is transcribed into RNA and is regulated by control regions.
Control regions are untranscribed sequences of DNA called enhancers and repressors, to which regulatory proteins produced by other genes bind.

- Enhancers and repressors sequences.
- Transcription factors bind to these sequences to promote or repress the transcription of DNA into a single stranded mRNA.

In eukaryotes, the transcribed sequence of a gene consists of coding regions called exons, separated by non-coding regions called introns.

After transcription is finished, the transcribed RNA sequence (pre-messenger RNA) is processed into an mRNA by splicing together the exons and removing the introns.

Alternative splicing takes when exons are spliced in different ways to produce different mRNA that can be translated into different proteins.

- Alternative splicing results in more proteins than genes in the genome.
- At least 35% of the human genes are subject to alternative splicing.

The genetic code is the sequence of nucleotides in DNA that specify for amino acids.

Triplets of nucleotides (codon) code for or specify all 20 the amino acids that make proteins.

There are 64 codons: 61 code for amino acids and 3 serve as stop signals.


Either RNA or DNA codons can be used to express the genetic code and amino acids the code for.

Sites: [http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/C/Codons.html](http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/C/Codons.html)

**CHARACTERISTICS OF THE GENETIC CODE**

- Messenger RNA consists of only four bases, A, G, C, and U, forming chains of various lengths and sequences.
- mRNA codon that specifies a given AA is a triplet of three nucleotides.
- Each codon is translated in a continuous sequence, three successive nucleotides at a time.
- The code is non-overlapping.
- The codon sequence complements an anticodon sequence in the adaptor tRNA.
- The code is universal. All living organisms share the codons that specify the same AA.
• The same codon does not specify two or more AA. There are no ambiguities.

• Except for methionine and tryptophan, all AA are designated by more than one codon. 64 codons specify 20 AA and chain termination.

• Degeneracy in the third codon position.

**Wobble hypothesis**: the third nucleotide of the tRNA anticodon can form hydrogen bonds with more than one kind of base in the third position of a codon.

• There are 64 codons, three of which are stop codons and bind to release factors rather than tRNA.
• The other 61 codons would need 61 tRNA but most organisms have less than 45 tRNA molecules.
• Therefore, the tRNA molecules should be able to bind to more than one codon.

Release factors are proteins that terminate translation by binding to a STOP codon.

The second position is less “degenerate”. A substitution of one base for another in the second position usually results in an amino acid substitution in the protein.

The universality of the genetic code is the basis for genetic engineering.

In eukaryotes, the vast majority of the DNA has no apparent function.

• Only about 1.5 to 2.5% of the human genome is thought to be transcribed. International Human Genome Sequencing Consortium (2001). "Initial sequencing and analysis of the human genome." *Nature* **409** (6822): 860-921.
• Recent research estimates 85 to 98% of the genome size to be non-coding. Much of this is comprised of repeat elements, transposons, and pseudogenes, but there is also a large amount of sequences that do not fall under any known classification.
• 45% of the human genome consist repeated sequences of 1 to 4 base pairs in length.
• These sequences are sometimes referred to as microsatellites.
• **Tandem repeats** occur when a pattern of nucleotides is repeated many times.
• Repeated sequences may be interspersed with other repeated sequences.

Microsatellites are a form of tandem repeats. Microsatellites are short sequences; tandem repeats are longer. Sometimes these terms are used interchangeably.

**Transposons** are segments of DNA that move and become inserted in different positions in the DNA.

• The process is called transposition.
• The moving sequence of DNA is called the transposable element, TEs.
• Transposition can duplicate the transposable element, TE, increasing the genome size.
• They are also called jumping genes.

Groups of genes that have similar sequences and related functions are called **gene families**.
Pseudogenes resemble functional genes but cannot be transcribed because they have in internal stop codon. They have lost the ability to code for proteins.

Processed pseudogenes originated by reverse transcription from an mRNA message into a DNA sequence that lacked introns, and later underwent silencing and further base pair changes.


**GENE MUTATIONS**

A gene is a functional unit.

Mutations are disruptions on the structure of a chromosome or changes in a single base pair of nucleotides.

Alleles are forms of a gene that produce a different phenotype.

Haplotypes refers to a collection of specific alleles of genes in a chromosome.

- Haplotypes is the contraction of “haploid genotype”.
- The sequence differs from its homologous sequence in the other chromosome.
- These alleles are closely linked and are less likely to be separated by recombination.
- The genes in a haplotype tend to be inherited together and not affected by recombination.

Genetic markers are sequences of DNA that can be used to identify regions of chromosomes or genes. They are differences between alleles of the same gene, e.g. single nucleotide polymorphism or SNP, a single base pair difference.

There are different types of genetic markers. Check the lists in these sites:

Genetic markers have a certain location in the gene or chromosome.

Mutations that occur in cells involved in reproduction are inherited.

Somatic mutations are not inherited unless the reproductive structures arise from somatic cells.

Helpful explanations:

**KINDS OF MUTATIONS**

**GENE or POINT MUTATIONS**
Point mutations are chemical changes in one base pair of a gene. They are also called SNPs (pronounced “snips”).

- **Base-pair substitution** in the DNA results in a different base pair that will be transcribed into an altered mRNA.
- **Transition** is a substitution of a purine for a purine, A↔G, or a pyrimidine for a pyrimidine, T↔C.
- **Transversions** are substitutions of purines for pyrimidines and vice versa.
- Because of the redundancy of the genetic code, **synonymous mutations** result in no changes in the amino acid sequence.
- **Non-synonymous** mutations result in the substitution of amino acids. They may or may not have consequences in the phenotype.

Insertions and deletions are additions or losses of nucleotide pairs in a gene.

In a **frameshift mutation** one or two nucleotides are inserted or deleted from the DNA.

As a result, the codons downstream of the insertion specify an entirely new sequence of amino acids.

The gene is greatly altered.

**REPLICATION SLIPPAGE**

**Replication slippage** alters the numbers of short repeats in microsatellites or coding sequences.

It can also occur in a coding region, which results in abnormal products.

DNA polymerase causes the slippage due to detachment and reattachment to the replicating strand.

**SEQUENCE CHANGES ARISING FROM RECOMBINATION.**

**Intragenic recombination:** An exchange of gene segments may occur between homologous sequences of DNA that differ at two or more base pairs thus forming a new combination DNA sequences.

In **gene conversion**, the gametes of a heterozygote should carry its two alleles in 1:1 ratio, but occasionally they occur in a different ratio, i.e. 1:3 ratio.

- One chromosome donates part of its genetic information to another chromosome.
- One chromosome is used as the template to correct a mismatched segment of another chromosome rather than its original strand.

“This conversion of one allele to the other is due to inappropriate base mismatch repair during recombination: if one of the four strands during meiosis pairs up with one of the four strands of a different chromosome, as can occur if there is sequence homology, mismatch repair can alter the sequence of one of the chromosomes, to match identically that on the other.”

http://en.wikipedia.org/wiki/Gene_conversion
http://www.web-books.com/MoBio/Free/Ch8D4.htm

**Unequal crossing over** can occur between two homologous sequences or chromosomes that are not perfectly aligned resulting in tandem duplication on one chromosome and deletion in the other.
CHANGES CAUSED BY TRANSPOSABLE ELEMENTS

1. **Insertion sequences**, which encode only enzymes that cause transposition.

2. **Transposable elements** produce copies that can move to any of many places in the genome, and sometimes they carry with them other genes near which they had been located.

   - They are called **transposons**.
   - The process is called **transposition**.
   - Transposons encode other functional genes.
   - They are also known as "jumping genes."

Transposons require **transposase** enzymes in order to be incorporated into a new location. This type of moving in the genome is called cut-and-paste transposition.

Several kinds of TEs include **insertion sequence**, which encode only enzymes that cause transposition.

3. **Retroelements** carry a gene for the enzyme **reverse transcriptase**.

   - Retroelements are first transcribed into RNA, which is then reverse-transcribe into a DNA copy, cDNA that is inserted in the genome.
   - Some retroelements are retroviruses that can move from cell to cell, e.g. HIV.
   - **Retrotransposons** are retroelements that cannot cross cell boundaries and are copied by cell division in the host.

Insertion sequences cause mutations when they land within a coding sequence of a gene, or a region of DNA that regulates gene expression.

   - May cause a frameshift
   - When inserted into a near a control region, they can interfere with or alter gene expression.
   - Increase mutation rates in host genes.
   - Can insert cDNA segments into RNA transcripts of other genes forming pseudogenes.
   - When a transposon leaves a gene, the resulting gap may not be repaired.
   - Many copies of transposons may cause unequal pairing resulting in unequal crossover.


**RATES OF MUTATIONS**

Mutation rate refers to the independent origin per gene copy per generation or per unit time.

Mutation rates are estimates, not absolutes, and these depend on the method used to detect mutations.

**ESTIMATING MUTATION RATES**

Rates can be estimated by

   - Counting the number of mutations arising in a laboratory stock, which is initially homozygous, scoring mutation effects either by their phenotypic effects or by molecular methods.
• **Base pair differences** between homologous genes in different species, relative to the number of generations that have elapsed since they diverged from their common ancestor.

Average mutation rate per base pair

• In prokaryotes is $10^{-10}$ per replication.
• In eukaryotes is $10^{-9}$ per gamete per generation.

The mutation rate in the human genome has been estimated at about $4.8 \times 10^{-9}$ per base pair per generation.

Mutation rates are low and not constant.

“In moderns humans carrying an estimated 25,000 genes per haploid genome, each sperm and egg may well carry less than one new mutation, or an average of $<0.4$ new mutations in a diploid fertilized zygote.” Strickberger’s Evolution. B. Hall and B. Hallgrimson. 2014. 5th Ed. Jones & Bartlett Learning, LLC, and Ascend Learning Company, USA. p. 368.

**Hot spots of mutations** occur when coiling of the DNA causes the replication errors by the polymerase enzymes. The nucleotide change may be imperfectly repaired or not repaired at all. This site may show mutations rates a hundred times higher than other sites.

**Back mutation** is mutation of a mutant allele back to the allele from which it arose.

• They usually occur at a much slower rate than forward mutations.
• At the molecular level, most are not restorations of the original sequence but instead result from a second amino acid substitution either in the same or different protein that restores the function that had been changed in the first substitution.

**EVOLUTIONARY IMPLICATIONS OF MUTATION RATES**

It has been estimated that the average human zygote will carry about 317 new mutations; if only 2.5% of the genome consists of transcribed sequences, 7 of these mutations will be expressed and will have the potential to affect the phenotype.

Mutation rates vary among genes and chromosome regions.

**Mutagens** are substances or other factors that cause mutation, e.g. UV light, many chemicals.

• Mutation rates are high for birds and mice in industrial areas.

The typical phenotypic character is **polygenic**, caused by several or many alleles in different loci.

**Mutational variance** is the increased variation in a population caused by new mutations in each generation.

The rate of origination of new genetic variation in the genome as a whole, and for individual polygenic characters is appreciable.

Mutation alone does not cause a character to evolve from one state to another.
PHENOTYPIC EFFECTS OF MUTATIONS

Mutations can affect fitness, e.g. reproduction and survival.

Phenotypic effects of mutational changes in DNA sequence range from none to drastic.

**Homeotic selector genes** determine the basic body plan of an organism and the sequence of developmental events conferring a distinct identity on each segment of the developing body by producing DNA-binding proteins.

- Mutation in these genes has grave consequences in the phenotype.
- They are usually highly conserved.
- For more information see the Principles course textbook by Campbell & Reece, chapter 21.

Related sites: [http://thebrain.mcgill.ca/flash/capsules/outil_rouge05.html](http://thebrain.mcgill.ca/flash/capsules/outil_rouge05.html)  

When two alleles of a pair are different, one is usually expressed and the other has no noticeable effect.

The expressed allele is called **dominant**, and unexpressed allele is called **recessive**.

A dominant allele can mask the expression of a recessive allele; this is called **complete dominance**.

**Incomplete dominance** occurs when the heterozygote has a phenotype intermediate between those of its parent.

- Red and white flower bearing plants produce a pink flower plant.

**Additive inheritance**: Characters expressed through the additive effect of many genes are called **quantitative characters**, e.g. human skin color.

EFFECTS OF MUTATIONS ON FITNESS

A few mutations slightly enhance fitness, some greatly decrease it and the majority has small deleterious effects.

The average effect of mutations that affects fitness is deleterious.

Many mutations are **pleiotropic**, that is, they affect more than one phenotypic trait.

- From the Greek *pleio*, "many", and *trepein*, "influencing or affecting".
- The gene codes for a product that operates as a signal on various tissues; the product has multiple molecular functions.
- Phenylketonuria in humans causes mental retardation and defective pigmentation of the skin and hair.
- The *yellow* mutation in Drosophila affects not only body color but also the male courtship behavior.
Mutations that are identified by their visible phenotypic effects often have deleterious pleiotropic effects.

Evolution would not occur unless some mutations were advantageous.

**THE LIMITS OF MUTATIONS**

Mutations alter pre-existing biochemical or developmental pathways.

Mutations cannot alter developmental foundations that do not exist.

- No mutation can make humans develop wings.

Some adaptive changes may not be possible without just the right mutation of just the right gene.

For these reasons, the rate and direction of evolution may in some instances be affected by the availability of mutations.

**Lethal mutations** represent changes in the genotype that are unable to interact successfully with what has evolved previously in the aspects of anatomy, physiology, etc.

What is a lethal change in one organism may not be lethal in another animal living in a different habitat. The processes of this second organism differ from the first one because of a different selective history.

It is the antecedent selective history and the intricately stabilized network of developmental genetic interactions that channel the subsequent evolution and keep the organism from developing novelties.

Compatibility of new features with established genetic functions serves as a powerful selective mechanism.

Selection that makes adaptive mutation for one situation makes those mutations un-adaptive for entirely different functions

- E.g. selection for excellence in swimming will conflict with traits that enable excellence in running.

**MUTATIONS AS A RANDOM PROCESS**

Mutations occur at random. This statement…

- Does not mean that all conceivable mutations are equally likely to occur, e.g. the developmental foundations do not exist.
- Does not mean that all loci or all regions within a locus are equally mutable: geneticists have described differences in mutation rates at both the phenotypic and molecular levels, among and within loci.
- Means that environmental factors do not induce mutations.

The probability of a mutation to occur does not depend on whether or not it would be advantageous.

Adaptive directed mutations do not occur. This is one of the fundamental tenets of the evolutionary theory.
RECOMBINATION AND VARIATION – *This section is not in your textbook*

All genetic variation owes its origin ultimately to mutation.

But in the short term, a great deal of genetic variation arises through recombination.

In sexually reproducing eukaryotes, genetic variation arises from two processes:

1. The union of genetically different gametes.
2. The formation of gametes with different combinations of alleles due to independent segregation of non-homologous chromosomes and to crossing over between homologous chromosomes.

Genotypes reassemble anew in each generation.

Recombination can slow down adaptation by breaking up favorable gene combinations, and enhance adaptation by providing natural selection with many combinations of alleles that have arisen by mutation.

---

ALTERATIONS OF THE KARYOTYPE

*Karyotype* refers to the number, size, shape and internal arrangement of the chromosomes of an individual, and to the photomicrograph showing the chromosomes.

*Nondisjunction* occurs when homologous chromosomes fail to segregate during meiosis.

*Aneuploidy* is the presence or absence of a single extra chromosome.

It is more common in humans than polyploidy.

- **Trisomic** individuals have three chromosomes of a kind.
- **Monosomic** individuals have a single chromosome of a kind.

Aneuploidy is due to abnormal mitosis or meiosis when the chromosomes fail to separate in the anaphase. This is called *nondisjunction*.

**POLYPLOIDY**

*Polyploidy* is the presence of several chromosome sets.

It is common in plants but rare in animals.

- Polyploidy is lethal in humans if it occurs in all cells of the body.
- A few triploid and tetraploid individuals have been born alive and survived a few days. They were found to have a mixture of diploid and polyploid cells.
- About 50–70% of angiosperms are thought to be polyploids.

Triploid gametes have aneuploid number chromosomes.
At segregation, each gamete will receive one copy of some chromosomes and two copies of others.

Euploidy refers to a balanced or complete number of chromosomes.

Autopolyploid is a polyploid formed from the doubling of a single genome. Polyploidy in which all the chromosomes come from the same species.

Allopolyploid is formed by hybridization of different but closely related species.

CHROMOSOME REARRANGEMENTS

Acrocentric chromosomes have the centromere near one end.

Metacentric chromosomes have the centromere somewhere in the middle and separated the chromosome into two arms.

Breakage of chromosomes can lead to four types of alterations in chromosomal structure.

» Deletions are loss of chromosomal material.
  - A chromosome breaks and fails to rejoin.
  - Most deletions are lethal.
  - Cri-du-chat is due to a deletion in chromosome 5.

» Duplication or fusion occurs when a piece of a chromosome breaks off, fission, and becomes attached on the sister chromatid causing a duplication of genetic material in the recipient chromosome.
  - Deletions and duplication are especially likely to occur during meiosis.

» An inversion happens when the detached piece is reattached to its chromosome but in the reverse orientation.
  - Inversions do not cause an imbalance in the genes but the change in location may influence the phenotype due its new location and neighboring genes, e.g. control regions.

» Translocation is the attachment of part of a chromosome to a nonhomologous chromosome.
  - Nonhomologous chromosomes may exchange parts, reciprocal translocation.
  - It may result in the elimination or duplication of genes.
  - A type of Down syndrome results from the translocation of a portion of chromosome 21 to chromosome 14. The individual has two normal chromosomes 21, one normal chromosome 14 and one abnormal chromosome 14 with a portion of chromosome 21 attached.
  - In chronic myelogenous leukemia (CML) a piece of chromosome 22 has switched places with a fragment from the tip of chromosome 9. The production of white blood cells is affected.
• Duplication and translocation tend to have harmful effects because essential genes may be affected.

Certain cancers are apparently due to chromosomal translocations,