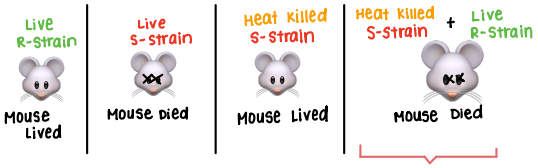
DNA Replication and Repair

- All the DNA in a cell constitutes the cell's genome
  - ↳ DNA molecules in a cell are packaged into chromosomes

- T.H. Morgan showed that genes are located on chromosomes
  - ↳ 2 components of chromosomes — DNA & protein — became candidates for the genetic material
    - ↳ Fredrick Griffith discovered the genetic role of DNA

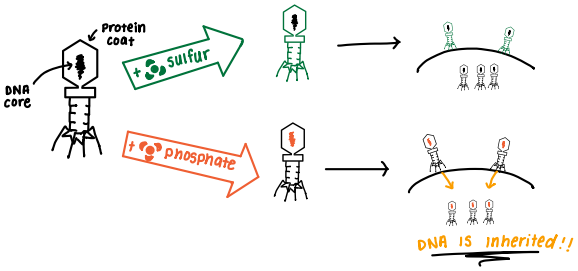
virulent S-Strain — Polysaccharide coating protects bacteria from the immune system

A-virulent R-Strain — No polysaccharide coating

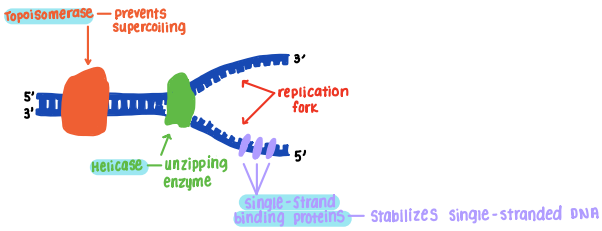


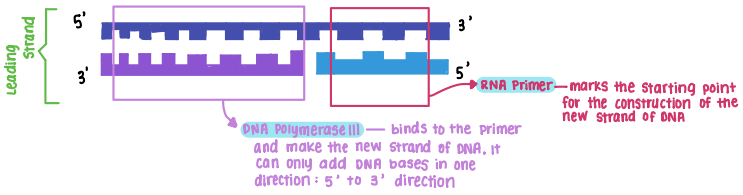
Individually, the heat killed S-Strain and the live R-Strain did not kill the mouse. So, something was transferred from heat killed S-Strain to live R-Strain that converted the live R-Strain to S-Strain

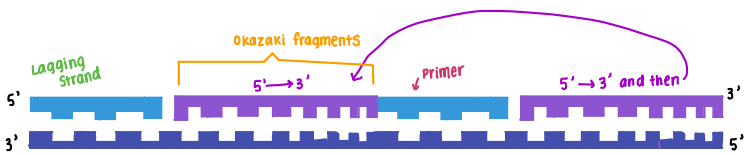
- Hershey-Chase Experiment : Is the genetic material DNA or Protein?



- 2 strands of DNA are complementary so each strand acts as a template for building a new strand in replication (semiconservative model)







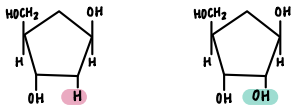
After formation of okazaki fragments, DNA Polymerase I removes the RNA primers & replaces the nucleotides with DNA. The remaining gaps are joined together by DNA ligase

Gene Expression: Transcription & Translation

DNA —transcription→ RNA —translation→ Protein

DNA vs. RNA

① DNA has **deoxyribose sugar** while RNA has **ribose sugar**



② RNA has the base **uracil (U)** while DNA has **thymine (T)**

A-U
C-G

A-T
C-G

② RNA is usually **single stranded & has a very short half-life** while DNA is **double stranded and quite stable**

Template Strand: 3'-ATGCGCCATGCCAGTTACGTAG-5'

↓ transcription

RNA Transcript: 5'-UACGCGGUACGGUCAAUGCAUC-3'

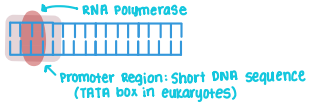
codons

During translation, the mRNA base triplets (codons), are **assembled in the 5' to 3' direction.**

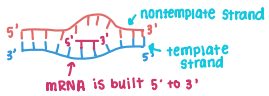
Transcription

- RNA is synthesized by **RNA polymerase**, which pries the DNA strands apart and assembles RNA nucleotides in the 5' to 3' direction **without the aid of a primer** (unlike DNA replication)
- Three stages:

↳ **Initiation:** Promoter is recognized, a bubble is created, and RNA synthesis begins

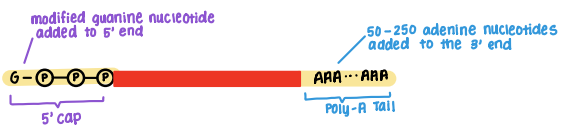


↳ **Elongation:** bubbles moves along the DNA template and transcript is elongated



↳ **Termination:** transcript and RNA polymerase are released and the bubble closes

- Enzymes in the eukaryotic nucleus modify pre-mRNA (RNA Processing) before the genetic messages are dispatched to the cytoplasm
  ↳ both ends of primary transcript are altered. Some interior parts of the molecule are cut out and the other parts spliced together.



- Most eukaryotic mRNAs have **long noncoding stretches of nucleotides (introns)** that lie between coding regions
  ↳ regions that are usually **translated into amino acid sequences are called exons.**
  ↳ RNA splicing removes introns and joins exons, creating an mRNA molecule with a continuous coding sequence.

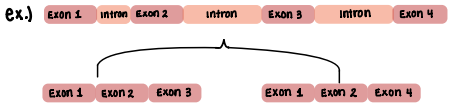
Pre mRNA:

introns cut out and exons spliced together

Mature mRNA:

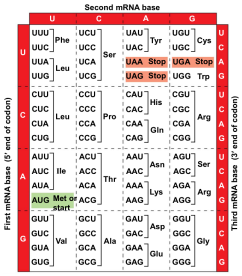
- Introns are removed in eukaryotes by a large complex called a **spliceosome** that is composed of different snRNPs ("snurps") which contains small nuclear RNA and a set of proteins

- **Alternative RNA splicing** occurs when many genes can give rise to 2 or more different polypeptides, depending on which segments are used as exons.



Translation

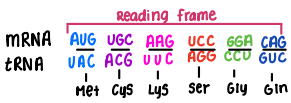
- In translation, the cell reads a genetic message (mRNA) and builds a polypeptide accordingly
- The translators are called transfer RNAs (tRNAs)
  ↳ tRNAs read the codons in mRNA & bring correct amino acids to the ribosome, which builds the polypeptide



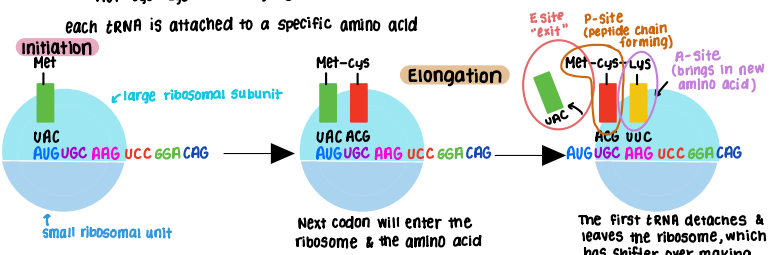
64 triplets; 61 code for amino acids & 3 triplets are stop codons

redundant: more than one codon may specify a particular amino acid
ex.) CGU & CGC both are Arg

not ambiguous: no codon specifies more than one amino acid
ex.) AGU = Ser NOT Arg



each tRNA is attached to a specific amino acid



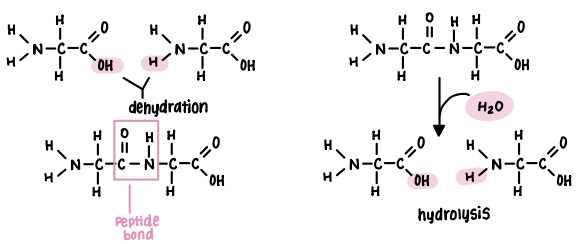
Small ribosomal subunit binds to an mRNA & initiator tRNA. Then the large ribosomal subunit joins to complete the translation initiation complex

Next codon will enter the ribosome & the amino acid will be covalently bound to Met

The first tRNA detaches & leaves the ribosome, which has shifted over, making room for the next tRNA

- During **elongation**, amino acids are added one by one to the previous amino acid at the C-terminus of the growing chain
  ↳ occurs in three steps:
  ① **Codon recognition** - codon binds to A-site
  ② **Peptide bond formation** - amino acid on P-Site binds to A-site
  ③ **Translocation** - A-site shifts to P-site

- **Termination** occurs when a stop codon in the mRNA reaches the A site of the ribosome. The A site accepts a protein called **release factor** that causes the addition of a water molecule instead of an amino acid. The polypeptide & ribosomal subunits and other components dissociate.
  ↳ hydrolysis



DNA Mutations & Repair Mechanisms

■ **Point mutations** are chemical changes in just one or few nucleotide pairs of a gene
↳ **nucleotide-pair substitution**: replace one nucleotide & its partner with another pair of nucleotides

- **Silent mutations** have no effect on the amino acid produced by a codon because of redundancy in the genetic code.
- **Missense mutations** still code for an amino acid, but not the correct amino acid.
- **Nonsense mutations** change amino acid codon into a stop codon, nearly always leading to a nonfunctional protein.

DNA Template strand: 3' TAC TTC AAA CCG ATT 5'
mRNA: 5' AUG AAG UUU GGC UAA 3'
Protein: Met – Lys – Phe – Gly – STOP

Silent mutation: 3' TAC TTC AAA CCA ATT 5' (A instead of G)
5' AUG AAG UUU GGU UAA 3' (U instead of C)
Met – Lys – Phe – Gly – STOP
→ no change in amino acids

Missense mutation: 3' TAC TTC AAA TCG ATT 5' (T instead of C)
5' AUG AAG UUU AGC UAA 3' (A instead of G)
Met – Lys – Phe – Ser – STOP
→ amino acid change to different amino acid

Nonsense mutation: 3' TAC ATC AAA CCG ATT 5' (A instead of T)
5' AUG UAG UUU GGC UAA 3' (U instead of A)
Met – STOP
→ premature termination of translation

■ **Frameshift mutation:** alter the reading frame of genetic message
↳ **insertions**—additions of nucleotide pairs in a gene
↳ **deletions**—losses of nucleotide pairs in a gene

Insertion: 3' TAC ATTC AAA CCG ATT 5' (Extra A)
5' AUG UAA GUU UGG CUA A 3' (Extra U)
Met – STOP
→ frameshift caused immediate nonsense

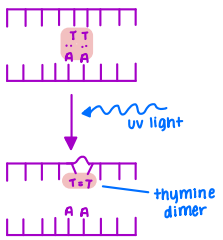
Deletion: 3' TAC TTC AACCG ATT 5' (A missing)
5' AUG AAG UUGGC UAA 3' (U missing)
Met – Lys – Leu – Ala …
→ frameshift caused extensive missense

Cause of Mutations

■ **Spontaneous mutations**—genetic changes that occur in the absence of mutagens & have no known cause
↳ **Errors in DNA replication**
↳ **Toxic metabolic products**: reactive chemicals such as free radicals
↳ **Changes in nucleotide structure**
↳ **Transposons**: small segments of DNA that can insert at various sites in the genome. If they insert into a gene, they may inactivate the gene.

■ **Induced mutations**—a mutation that arises from exposure of an organism's DNA to a mutagen (physical or chemical)
↳ higher rate than spontaneous mutations
↳ **chemical agents**: chemical substances, such as benzo(a) pyrene, a chemical found in cigarette smoke
↳ **physical agents**: UV (ultraviolet) light & X-rays

■ **Physical Mutagens**
↳ **ionizing radiations** (ex.) x-rays & gamma rays)—cause deletions or breaks in one or both DNA strands
↳ **Nonionizing radiations** (ex.) UV light



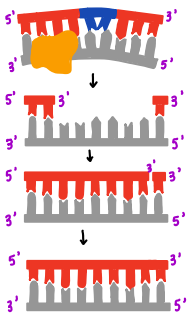
UV light & some chemicals can cause thymine dimers to form. These dimers produce a kink in the DNA strand and block DNA replication

Repair of Mutations

■ Two components: **Detection** of damage & **Repair** of damage

■ DNA replication is very accurate. DNA polymerase matches bases with high accuracy. It inserts an incorrect base about once every 100,000 bases
↳ **Repair enzymes remove defective bases and replace them with the correct one.**
↳ DNA polymerase can proofread its work. DNA polymerase will add a nucleotide only if the previous base pair is correct
- If the enzyme finds a mismatch, it removes the mismatched base

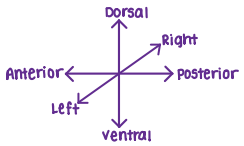
■ The nucleotide excision repair (NER) system recognizes different types of damage:



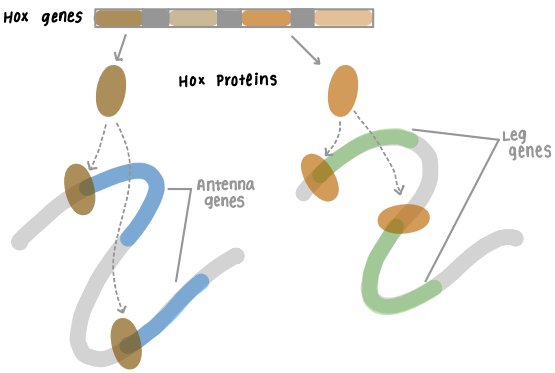
UvrA-UvrB complex tracks along the DNA in search of damaged DNA. Removes the damaged single-stranded DNA (tymine dimer in this example). Uses the intact strand a template for new DNA. DNA ligase links the repaired strand to the original undamaged DNA.

Gene Regulation & Development in Eukaryotes

- **Programmed cell death** is highly regulated and essential part of development
  - ↳ occurs in both plants and animals
  - ↳ Helps tissues and organs take shape
  - ↳ **Apoptosis** is the most common type in animals
    - cells that form webbing between toes die
    - About half of neurons die as nervous system is wired
    - Harmful immune cells are eliminated
- **Pattern formation** is the development of a spatial organization of tissues and organs. Begins in the early embryo, in animals, with the establishment of the major axes. **Positional information** is the molecular cues that control pattern formation, tells a cell its location relative to the body axes and to neighboring cells



- **Maternal effect genes** — mom's genes effect phenotype
  - ↳ nurse cells: dumps mRNA in egg before fertilization (effects orientation)
  - ↳ bicoid gene affects the front half of the body
    - An embryo whose mother has no functional bicoid gene lacks the front half of its body & has duplicate posterior structures at both ends
    - Bicoid mRNA is highly concentrated at the anterior end. After the egg is fertilized, the mRNA is translated into Bicoid protein, which diffuses from the anterior end (a gradient of Bicoid protein)
- Lewis discovered the **homeobox (Hox) genes**: master control (homeotic) genes that initiate pattern formation (organs development in specific parts of the body) in late embryo, larva, and adult stages.
  - ↳ Homebox (Hox) genes guide pattern formation in embryos — tell the cells of the body how to differentiate as the body grows.
- **Homeosis** is the transformation of one organ into another via mutation. Homeosis is caused by mutations in Hox genes.



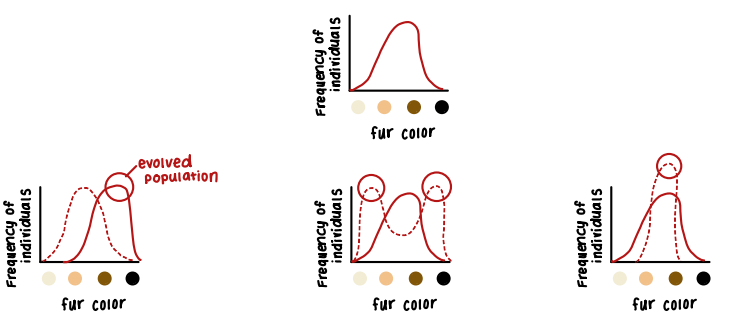
- Hox genes each encode a 180 nucleotide stretch that is translated into a homeodomain in the protein; homeodomains bind different DNA sequences. Hox proteins are transcription factors that regulate many other genes.

Microevolution & Population Genetics

- **Microevolution** is a change in allele frequencies in a population over generations
  - ↳ THREE mechanisms cause allele frequency change
    - Natural selection
    - Genetic Drift (chance events altering allele frequency)
    - Gene flow (transfer of alleles between populations)
  - ↳ only **natural selection causes adaptive evolution** (improving the match between organisms and their environment)
- variation in heritable traits is a prerequisite for evolution. **Without genetic variation, evolution CANNOT occur.**
- **New genes and alleles can arise by mutation or gene duplication. Only mutations in gametes can be passed to offspring.**
  - ↳ **the effects of point mutations can vary**
    1. Mutations in noncoding regions of DNA are often harmless
    2. Mutations to genes can be neutral because of redundancy in genetic code
    3. Mutations that alter the phenotype are often harmful
    4. Mutations that result in a change in protein production can sometimes be beneficial
  - ↳ **duplicated genes can take on new functions by further mutation**
    - An ancestral odor-detecting gene has been duplicated many times: Humans have 350 functional copies of the gene
- A **population** is a localized group of individuals capable of interbreeding and producing fertile offspring.
- A **gene pool** consists of all the alleles for all loci in a population.
- **The Hardy-Weinberg Principle** states that frequencies of alleles and genotypes in a population remain constant from generation to generation. The Hardy-Weinberg principle describes a population that is NOT evolving.
  - ↳ The frequency of an allele in a population can be calculated
    - By convention, if there are 2 alleles at a locus, p and q are used to represent their frequencies (example: p represents A; q represents a)
    - **The frequency of all alleles in a population will add up to 1.**
- For example, consider a population of wildflowers that is incompletely dominant for color
  - 320 red flowers (C^RC^R)
  - 160 pink flowers (C^RC^W)
  - 20 white flowers (C^WC^W)
  - ↳ calculate the number of copies of each allele
    - $C^R = 800$
    - $C^W = 200$
  - ↳ calculate the frequency of each allele
    - $C^R = \frac{800}{(800+200)} = 0.8\ (80\%)$
    - $C^W = 1 - 0.8 = 0.2\ (20\%)$
  - ↳ The sum of alleles is always 1
- **Hardy-Weinberg equilibrium** describes the constant frequency of alleles in such a gene pool.
  - ↳ From the example above, if p (C^R) and q (C^W) represent the frequencies of the only two possible alleles in a population at a particular locus, then 3 genotypes will occur in the following proportions:

$$p^2 + 2pq + q^2 = 1$$

  - ↳ The frequency of genotypes can be calculated
    - $C^RC^R = p^2 = (0.8)^2 = 0.64 = 64\%$
    - $C^RC^W = 2pq = 2(0.8)(0.2) = 0.32 = 32\%$
    - $C^WC^W = q^2 = (0.2)^2 = 0.04 = 4\%$
- The **Hardy-Weinberg theorem** describes a hypothetical population that is NOT evolving
- The **five conditions for non-evolving populations** are rarely met in nature
  1. No mutations
  2. Random mating (no inbreeding)
  3. No natural selection (no differences in survival or reproductive success)
  4. Extremely large population size (small population—more bias allele)
  5. NO gene flow
- Three modes of natural selection:



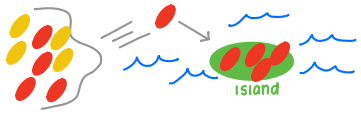
Directional selection— favors individuals at one end of the phenotypic range

Disruptive selection— favors individuals at both extremes of phenotypic range

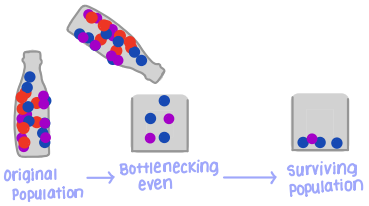
Stabilizing selection— favors intermediate variants and acts against extreme phenotypes.

- **Genetic Drift** describes how **allele frequencies fluctuate unpredictably from one generation to the next.** The smaller a sample or population, the more likely it is that chance alone will cause deviation from a predicted result. **"Chance" events affect survival and reproduction.** Genetic drift tends to reduce genetic variation through losses of alleles, especially in small populations.
  - ↳ **Subtypes: Founder Effect and Bottleneck Effect**

- **The founder effect** occurs when a few individuals become **isolated from a larger population.**
  - ↳ Allele frequencies in the small founder population can be different from those in the larger parent population due to chance.



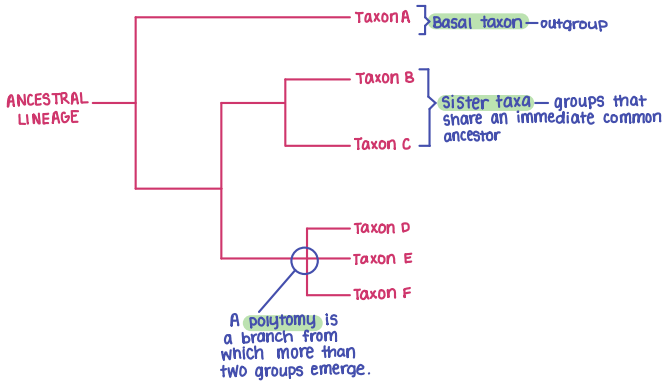
- **The bottleneck effect** can result from a drastic reduction in population size due to a sudden environmental change.
  - ↳ By chance, the resulting gene pool may no longer be reflective of the original population's gene pool.
  - ↳ If the population remains small, it may be further affected by genetic drift.
  - ↳ If the population recovers, it may not be very genetically diverse.



- Genetic drift is significant in small populations; alleles can be disproportionately under- or over-represented in the next generation.
- **Gene flow** consists of the movement of alleles among populations
  - ↳ Alleles can be transferred through the movement of fertile individuals or gametes (ex: pollen)
  - ↳ Gene flow tends to reduce genetic variation among populations over time.
- Genetic drift and gene flow do not consistently lead to adaptive evolution, as they increase or decrease the match between an organism and its environment.

Microevolution & Phylogeny

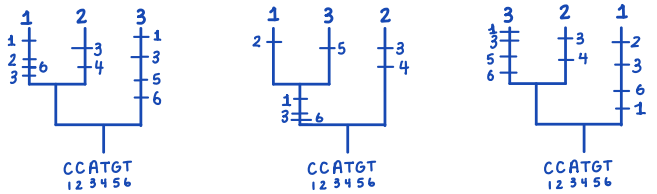
- **Macroevolution**: bigger changes that happen above the level of species (ex: the origin of mammals or radiation of plants
- **Speciation** is the process by which one species splits into two or more species
- **Phylogeny** is the evolutionary history of a species or group of related species
- **Taxonomy** is the ordered division and naming of organisms
- The two-part scientific name (Latin) of a species is called a **binomial**
  - ↳ The first part of the name is the **genus**
  - ↳ The second part, called the *specific* epithet, is unique for each **species** within the genus.



- Phenotypic and genetic similarities due to shared ancestry are called **homologies**.
  - ↳ organisms with similar morphologies or DNA/protein sequences are likely to be more closely related than organisms with different structures or sequences
- Phylogenic trees show patterns of descent, not phenotypic similarity
- Phylogenic trees do not generally indicate when a species evolved or how much change occured in a lineage.
- It should not be assumed that a taxon evolved from the taxon next to it (ex: wolves didn't evolve from coyotes; they both evolved from a common ancestor that was neither wolf or coyote and no longer exists.)
- **Homology** is similarities due to shared ancestry (close common ancestor)
- **Analogy** is similarities due to convergent evolution (look the same but different ancestor)
  - ↳ **convergent evolution** occurs when similar environmental pressures and natural selection produce similar (analogous) adaptations in organisms from different evolutionary lineages.
  - ↳ Analogous structures or molecular sequences that evolved independently are also called **homoplasies**.
- A **clade** is a group of species that includes an ancestral species and all its descendants
- The principle of **maximum parsimony**: the tree that requires the fewest evolutionary events

example:

	1	2	3	4	5	6
Ancestral species:	C	C	A	T	G	T
Species #1:	A	T	C	T	G	C
Species #2:	C	C	T	G	G	T
Species #3:	A	C	C	T	A	C



Tree A
10 nucleotide changes

Tree B
7 nucleotide changes

Tree C
10 nucleotide changes

Parsimonius Tree: Tree B