2027 Heredity/Designer Genes - Introduction and Learning Objectives

How to use the Heredity/Designer Genes Resources

- If you are just getting started learning about genetics, we suggest you carefully read through this document. This will help you understand the basic principles of genetics and provide a foundation for further study.
- Once you feel you have a strong understanding of the foundational principles covered here, dive into specific topics! We provide suggested resources specific to each topic. These are suggestions - there are many genetics resources available online and we encourage you to explore and find what works best for your learning style.
- The recommended textbooks can be used as a starting point for studying and as a tool in testing your knowledge
    - Most textbooks have end-of-unit questions to check understanding
    - Suggested exercise: find textbook figures relevant to the topic you are studying and write a caption for the figure that describes what the figure is communicating.
- Test your knowledge by answering practice questions - see the Practice Tests provided on the resource page (keep in mind these tests are written based on previous years' rules); write practice questions for your study partner to answer and vice versa!

General Resources

There are many excellent online and printed resources that cover a wide breadth of topics in microbiology.

- Online Open Genetics Textbook - Provides a great introduction to the topics covered in the event, particularly recommended for **Division B competitors.**
- Genetics, from Genes to Genomes, an excellent textbook used in many undergraduate genetics courses. Previous additions (5th, 6th,7th) may be found online for lower prices. This book covers topics in great depth - be sure to cross reference your reading with the study guides to make sure you don't accidentally spend time on topics beyond the scope of the event. This textbook is recommended for **Division C competitors**.
- Any relatively recent (c. 2015 or later) genetics textbook you can get your hands on! Check any libraries you have access to, including your school, town, or nearby colleges and used bookstores.
- University of Utah Genetic Science Learning Center - Educational modules and interactive activities on many of the topic areas covered by the Heredity/Designer genes rules

Topic-Specific Resources

Mendelian Genetics

- Khan Academy Classical Genetics: https://bio.libretexts.org/Bookshelves/Genetics/Classical_Genetics_(Khan_Academy)

Mitosis and Meiosis

- How cells divide: https://www.pbs.org/wgbh/nova/miracle/divide.html

Population & Evolutionary Genetics

- Population genetics explorer: https://media.hhmi.org/biointeractive/click/population-genetics-explorer/index
- Virtual Biology lab on Population genetics: https://virtualbiologylab.org/population-genetics/
- Virtual Biology lab on Selection: https://virtualbiologylab.org/selection/
- Creating phylogenetic trees from DNA sequences: https://www.biointeractive.org/classroom-resources/creating-phylogenetic-trees-dna-sequences

Molecular Biology of DNA

- Working with Molecular Genetics: https://bio.libretexts.org/Bookshelves/Genetics/Working_with_Molecular_Genetics_(Hardison)
- Genetic mutations and disease: https://www.biointeractive.org/classroom-resources/genetic-mutations-and-disease-0

Prokaryotic gene expression and regulation

- Working with Molecular Genetics: https://bio.libretexts.org/Bookshelves/Genetics/Working_with_Molecular_Genetics_(Hardison)
- Lac operon simulation: https://www.labxchange.org/library/items/lb:LabXchange:103ca5bd:lx_simulation:1
- Trp operon simulation: https://www.labxchange.org/library/items/lb:LabXchange:e8c2344c:lx_simulation:1?_gl=1*bqs3lc*_up*MQ..*_ga*MTQ5NDg0MjM4My4xNzU2NDA1MTY3*_ga_JPPNBMEJG0*czE3NTY0MDUxNjYkbzEkZzAkdDE3NTY0MDUxNjYkajYwJGwwJGgw

Technology and Techniques

- PCR: https://www.labxchange.org/library/items/lb:LabXchange:e145d666:lx_simulation:1?_gl=1*1onz1o4*_up*MQ..*_ga*MTQ5NDg0MjM4My4xNzU2NDA1MTY3*_ga_JPPNBMEJG0*czE3NTY0MDUxNjYkbzEkZzEkdDE3NTY0MDUxNzUkajUxJGwwJGgw
- Sanger sequencing video: https://www.biointeractive.org/classroom-resources/sanger-sequencing
- Introduction to Genetic Engineering: The Process https://www.labxchange.org/library/pathway/lx-pathway:b0fd731a-5bab-4fe2-8491-405292e5176e?source=%2Flibrary%2Fclusters%2Flx-cluster%3Aabe&_gl=1*1quz2mr*_up*MQ..*_ga*ODYxNTU4NzcyLjE3NTY0MDYwNjA.*_ga_JPPNBMEJG0*czE3NTY0MDYwNTkkbzEkZzEkdDE3NTY0MDYwNzAkajQ5JGwwJGgw
- Building a Recombinant Plasmid: Restriction Enzymes https://www.labxchange.org/library/pathway/lx-pathway:27bdb1d1-b7cf-4d7d-95fb-e28e9c6e21f9
- CRISPR-Cas9 interactive: https://www.biointeractive.org/classroom-resources/crispr-cas9-mechanism-applications

PRINCIPLES OF GENETICS - Resources Developed by Karen Lancour

This section should <u>not</u> be interpreted as an extension of the rules

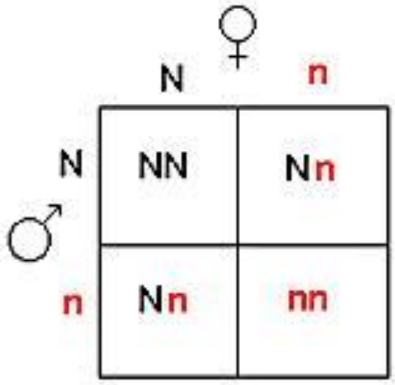
Basic Principles

gene - a unit of inheritance that usually is directly responsible for one trait or character. Each individual has two genes for each trait, one comes from dad and the other from mom.
allele - alternate forms of a gene. Usually there are two alleles for every gene, sometimes there are more than two alleles present in population – termed **multiple alleles**
homozygous - when the two alleles are the same
heterozygous - when the two alleles are different

dominant - a trait (allele) that is expressed irregardless of the second allele
recessive - a trait that is only expressed when the second allele is the same (e.g. short plants are homozygous for the recessive allele)

- **Dominant –** always expressed
  - **Capital letters – N**
  - **Homozygous - NN**
  - **Heterozygous - Nn**
- **Recessive –** prevented by dominant
  - **Lower case letters – n**
  - **Homozygous - nn**

Punnett square - probability diagram illustrating the possible offspring of a mating male genes on top of columns and female traits on side of rows



Dominant and Recessive

Autosomal Dominant - dominant gene on an autosome
Autosomal Recessive - recessive gene on an autosome
Sex-linked Dominant - dominant gene on a sex chromosome
Sex-linked Recessive – recessive gene on a sex chromosome

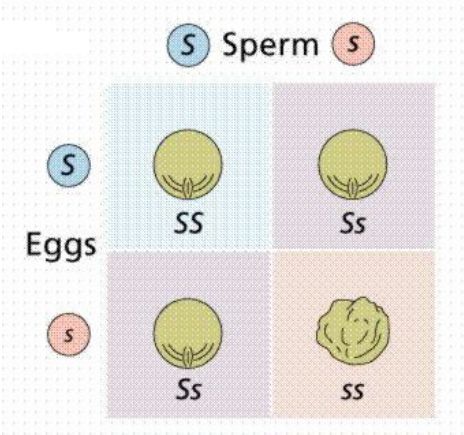
Genotype and Phenotype

phenotype - the physical expression of the genes for the trait by an individual
genotype - the gene makeup of an organism. Phenotype is the trait of an individual expresses while genotype is the two genes that cause that trait

Monohybrid Cross – a cross involving only one trait.

(phenotype ratio – 3:1 and genotype ratio 1:2:1)

hybrid – an individual who has one dominant and one recessive gene for a trait

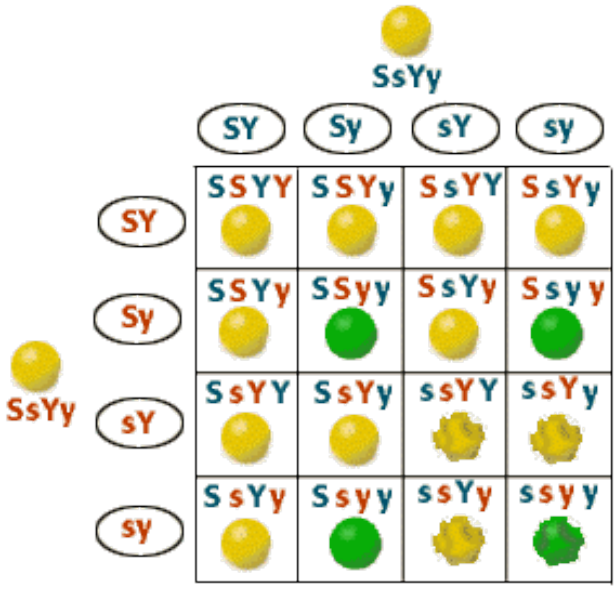


- **Hybrid – Ss X Ss**
- **One Trait – Smooth vs wrinkled**
- **Two gametes per parent**
- **S and s**
- **Punnett Square with 4 boxes – 4 offspring**

Genotype	**Phenotype**
SS or Ss	Round
ss	Wrinkled

Genotype frequency
1:2:1

Phenotype frequency
3:1



Dihybrid Cross – a cross involving two traits.

(phenotype ratio-9:3:3:1 and genotype ratio- 1:2:1:2:4:2:1:2:1)

Dihybrid – 2 traits
Gametes per parent = 4
Punnett Square – 16 boxes
Genotype ratio
1:2:1:2:4:2:1:2:1
Phenotype ratio
9:3:3:1

Trihybrid Cross

triple-het x triple-het cross

SsYyAa x SsYyAa

	SYA	SYa	SyA	Sya	sYA	sYa	syA	sya
SYA	SSYYAA	SSYYAa	SSYyAA	SSYyAa	SsYYAA	SsYYAa	SsYyAA	SsYyAa
SYa	SSYYaA	SSYYaa	SSYyaA	SSYyaa	SsYYaA	SsYYaa	SsYyAa	SsYyaa
SyA	SSyYAA	SSyYAa	SSyyAA	SSyyAa	SsyYAA	SsyYAa	SsyyAA	SsyyAa
Sya	SSyYaA	SSyYaa	SSyyaA	SSyyaa	SsyYaA	SsyYaa	SsyyaA	Ssyyaa
sYA	sSYYAA	sSYYAa	sSYyAA	sSYyAa	ssYYAA	ssYYAa	ssYyAA	ssYyAa
sYa	sSYYaA	sSYYaa	sSYyaA	sSYyaa	ssYYaA	ssYYaa	ssYyaA	ssYyaa
syA	sSyYAA	sSyYAa	sSyyAA	sSyyAa	ssyYAA	ssyYAa	ssyyAA	ssyyAa
sya	sSyYaA	sSyYaa	sSyyaA	sSyyaa	ssyYaA	ssyYaa	ssyyaA	ssyyaa

Phenotypes:
Out of 64 births.

- normal
- albinos
- anerythristic
- striped
- snow
- striped-albino
- striped-anery
- striped-snow

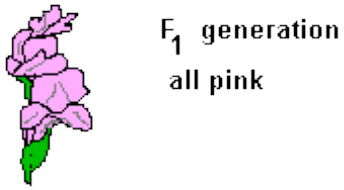
Incomplete Dominance

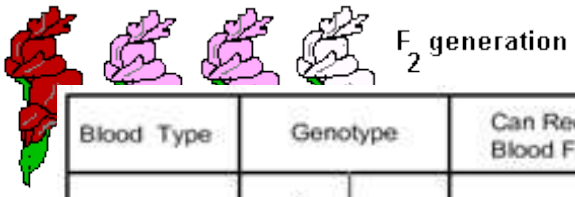
Incomplete dominance – one allele (gene) is not completely dominant over another resulting in a blending of traits and where the phenotype of a hybrid displays a blending of the two alleles

Incomplete dominance –

- **Hybrid** is a blend of two traits
- **Genotype frequency**
  1:2:1
- **Phenotype frequency**
  1:2:1
- **Examples:** Flowers, Animal fur







Co-dominance

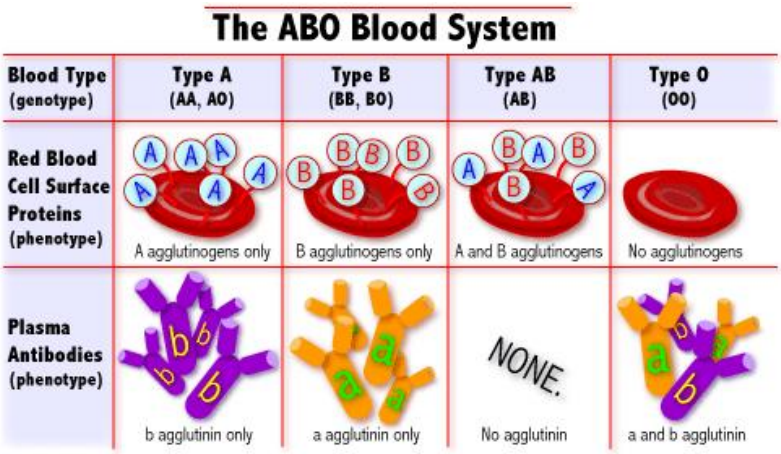
co-dominance – both dominant alleles (genes) in an individual are expressed as in blood types

Blood types – A,B,O alleles

A and B genes are co-dominant and both dominant over the O gene which is recessive

Phenotypes	Genotype
A	**I^AI^A or I^Ai**
B	**I^BI^B or I^Bi**
AB	**I^AI^B**
O	**ii**

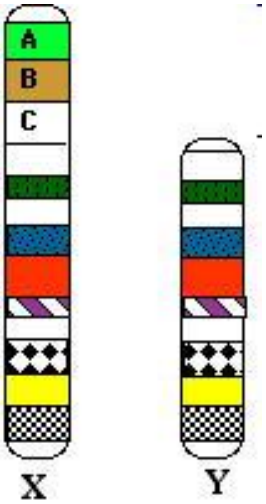
Blood Type	Genotype		Can Receive Blood From:
A	i^Ai i^Ai^A	AA AO	A or O
B	i^Bi i^Bi^B	BB BO	B or O
AB	i^Ai^B	AB	A, B, AB, O
O	ii	OO	O



Independent Assortment vs. Linkage

- **Independent Assortment** – genes on different chromosomes separate independently during meiosis
- **Linkage –** genes on the same chromosome are inherited as a group
    - **Autosomal linkage –** on autosomes
    - **Sex-linked –** on sex chromosomes

Linkage – Sex Linkage



- **Linkage –** genes on the same chromosome inherited as a group
- **Sex-linkage –** genes on sex chromosomes (esp. X)
- **Y-chromosome shorter –** some genes from X missing
- X-linked traits more common in men
- Men get X-chromosome from mom
- Red-green colorblindness, hemophilia

sex-linkage – allele (gene) is located on a sex chromosome and it will be more common in one sex.

- It is usually on the x-chromosome and more common in males than in females.
- Barr bodies – tightly coiled X chromosome in females – inactive X chromosome.
- Calico cats – usually on females. yellow and black alleles on X chromosome - female has 2 X's

Genetic Variations

Probability – ratios or percents
Multiple Alleles – three or more alleles for a gene as blood type as skin color
Multifactorial Traits – more than 1 pair of genes plus environment
Pleiotropy – the action of an allele (gene) affects many parts of the body as sickle cell anemia
Variable expressivity – an allele (gene) can be expressed differently in different people

Environmental influence on genes expression

- Gene function is influenced by environment as with identical twins
- Genes have blueprint for proteins or parts of proteins
- Proteins can be structural proteins (parts of body) or functional proteins (hormones/enzymes

Epistasis and Multifactorial Inheritance

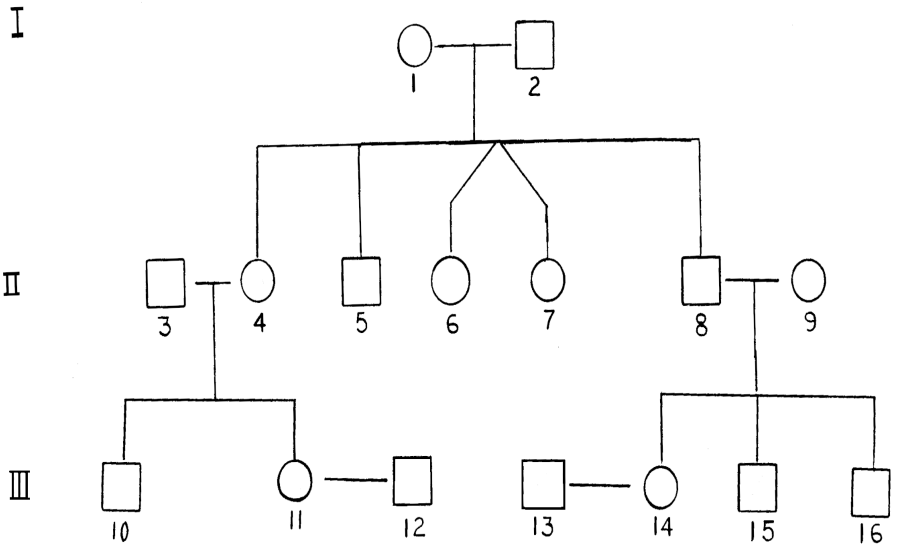
- **Epistasis -** the interaction between two or more genes to control a single phenotype so one pair of genes alters the expression of another pair of genes as albino
- **Multifactorial inheritance -** many factors (multifactorial) both genetic and environmental are involved in producing the trait or condition. Examples: height, weight, cleft palate, spina bifida

Pedigree Analysis

Pedigree is a family tree**.**

Squares are males and circles are females

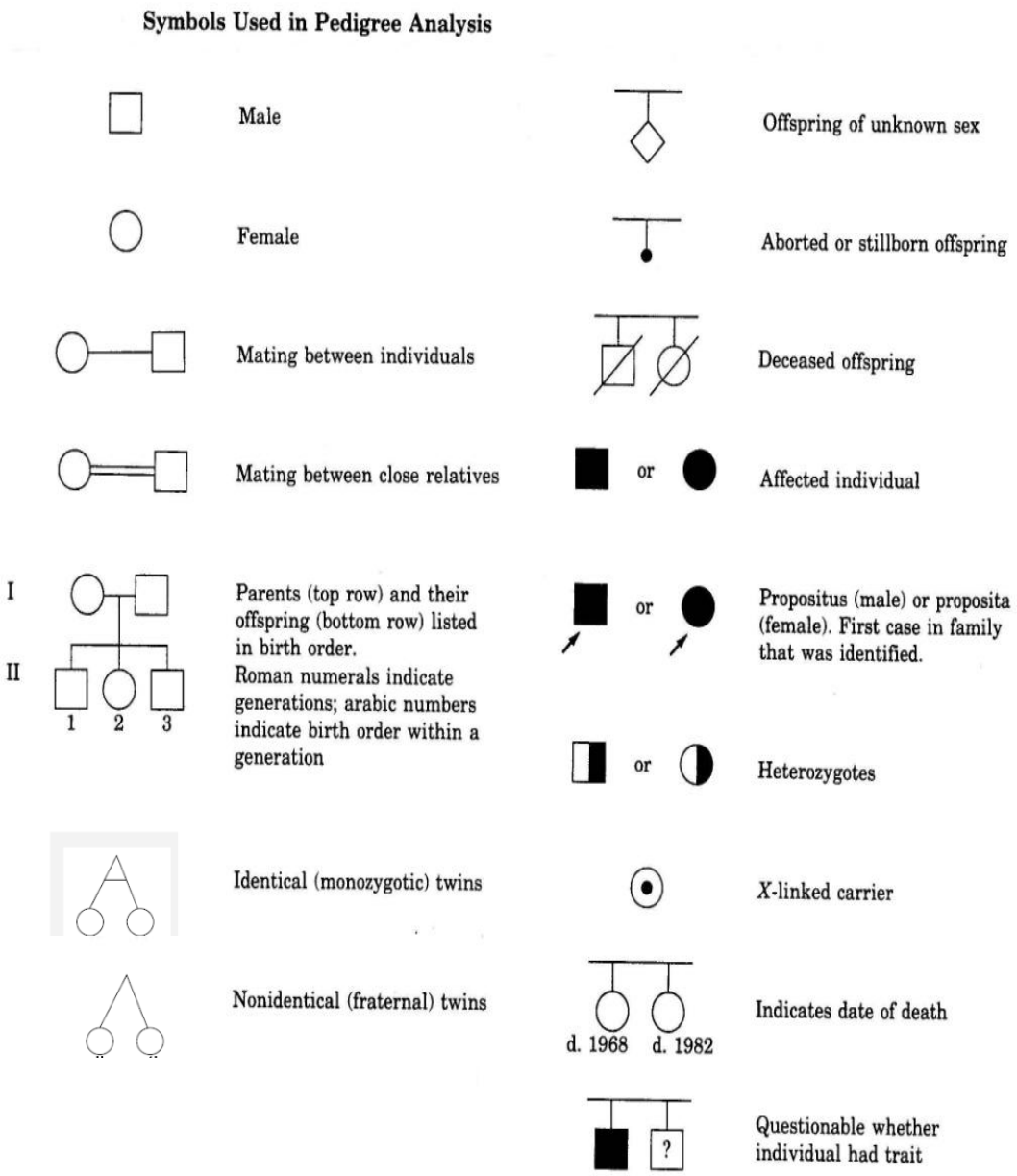
Generations = I – Original Parents
II- F1 (children)
III – F2 (grandchildren)



Karyotype Analysis - **karyotype** is print of human chromosomes

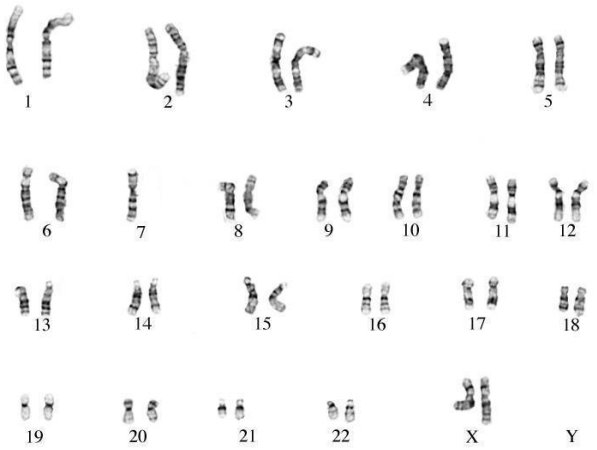
- **nondisjunction** – chromosomes do not separate during meiosis.
- Results in monosomy and trisomy

Karyotype Characteristics:



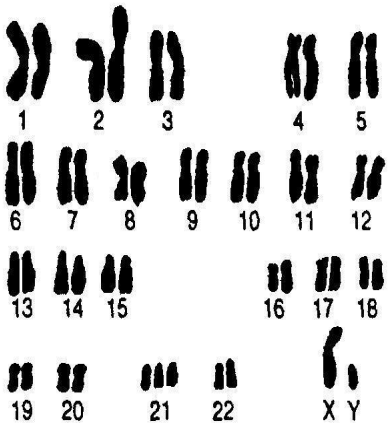
- The numbered chromosome pairs termed autosomes are arranged longest to shortest

- Chromosomes come in pairs
- The sex (X & Y) chromosomes are placed last with normal females having XX and normal males having XY
- If only X chromosomes are present, it will be female
- If X and Y chromosomes are present, it will be male
- Bent chromosomes are not abnormal. It is just the way they were photographed
- If an individual has an extra chromosome, it is termed **trisomy** and if a chromosome is missing, it is termed **monosomy**



Karyotype 1

Male with Monosomy 7



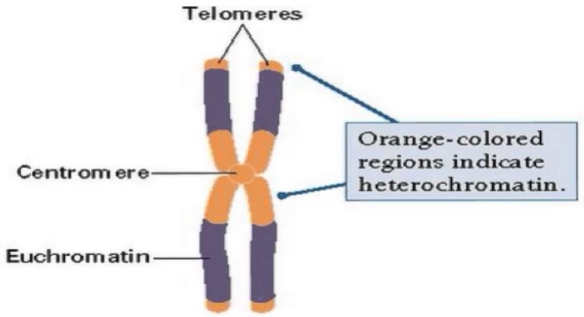
Karyotype 2

Female with Trisomy 21

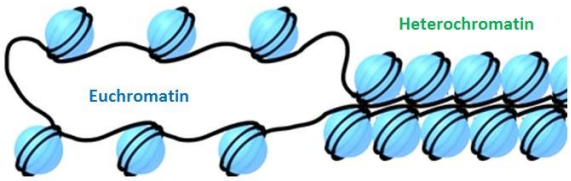
Regions of Chromosome for Replication

Heterochromatin vs Euchromatin

Heterochromatin DNA is generally not very active
In interphase this type of chromatin are compact
Heterochromatin also has long stretches of repeat sequences called satellite DNA
Euchromatin is generally more active
In interphase they are generally not condensed

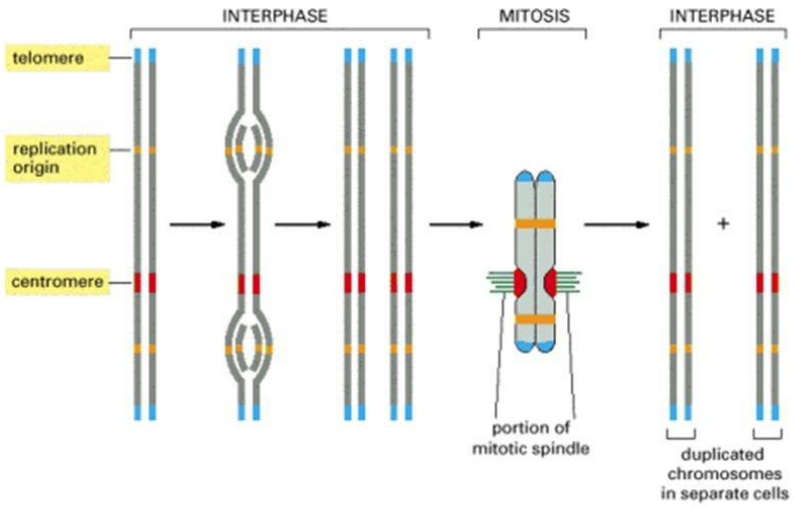


Difference between heterochromatin and euchromatin



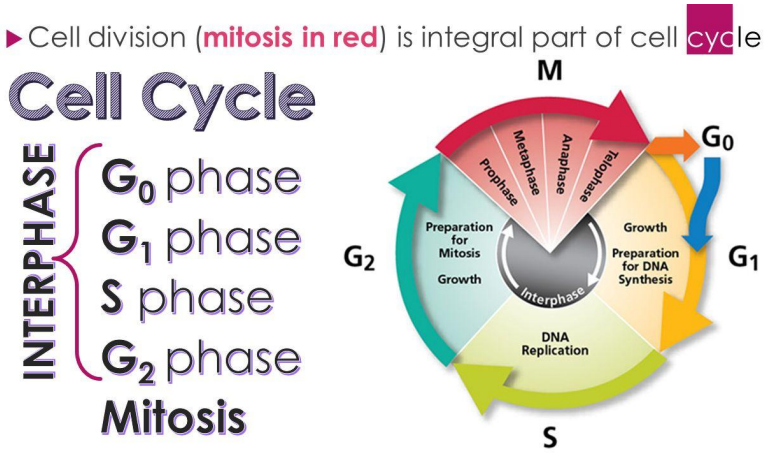
- **Heterochromatin** is a part of chromosome, a tightly packed form of DNA whereas **euchromatin** is an uncoiled form of chromatin.
- **Heterochromatin** has tighter DNA packing than **euchromatin**.
- **Heterochromatin** stains dark in interphase whereas **euchromatin** stains lightly with basic dyes but stains dark during mitosis, when it is in condensed state during each repetition of the cell cycle.
- **Heterochromatin** contains more number of DNA compare to **euchromatin**.
- **Heterochromatin** found in eukaryotes whereas **euchromatin** found in both eukaryotes and prokaryotes.
- **Heterochromatin** is genetically inactive and **euchromatin** is genetically active.
- **Heterochromatin** is late replicative whereas **euchromatin** is early replicative.

Each DNA Molecule That Forms a Linear Chromosome Must Contain a **Centromere**, Two **Telomeres**, and **Replication Origins**



MITOSIS, MEIOSIS, ASEXUAL VS. SEXUAL REPRODUCTION

Cell Cycle – the life cycle of a cell



G_0 *Phase*– Cells that go into this phase when not actively reproducing as muscle or nerve cells

G_1 ***Phase*** **–** high rate of biosynthesis and growth

S *Phase* **–** DNA content doubles and chromosomes replicate

G_2 ***Phase*** **-** final preparations for Mitosis

M ***Phase*** **–** Mitosis and Cytokinesis

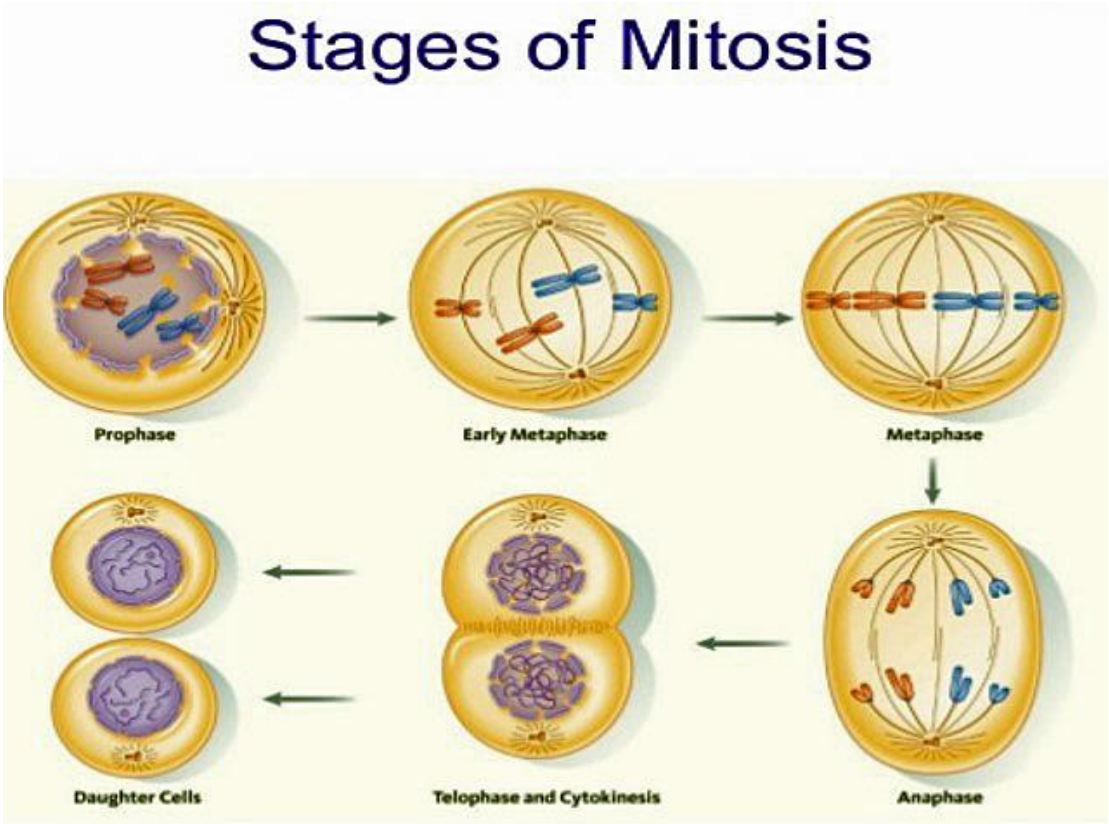
A. ***Prophase*** **–** chromatid pairs coil up, spindle forms, nuclear membrane dissolves, chromatid pairs attach to spindle fibers (microtubules),

B. ***Metaphase*** **–** chromatid pairs move to the equator, chromatid pairs align at the equator,

C. ***Anaphase*** **–** chromatids separate into individual chromosomes, chromosomes are pulled apart toward the equator by the spindle fibers (microtubules)

D. ***Telophase*** **-** chromosomes uncoil, spindle dissolves, nuclear membrane reforms

E. Cytokinesis **–** division of the cytoplasm to make two new cells



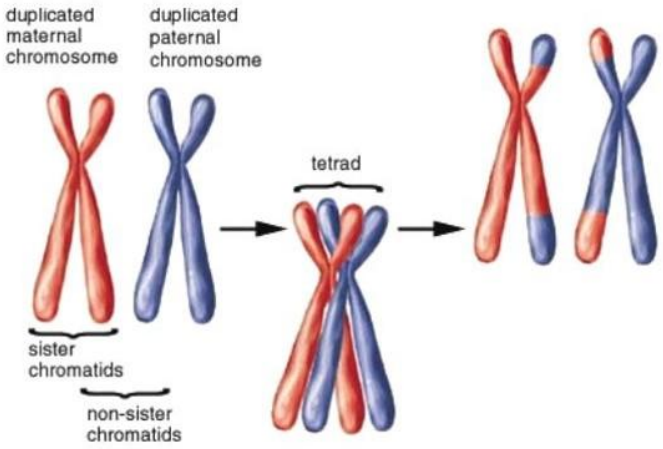
Mitosis vs Meiosis

Mitosis type of cell reproduction which produces two daughter cells that are genetically identical to the parent cell.

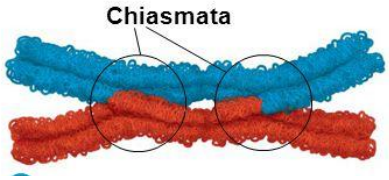
- ❖ **Growth and Asexual Reproduction**
- ❖ **One division – 2 diploid cells**
- ❖ **Genetically same as original**

- **Meiosis** type of cell reproduction which produces 4 cells which half the number of chromosomes as the original parent cell
  - ❖ **Gametes for Sexual Reproduction**
  - ❖ **2 divisions – 4 haploid cells**
- **Sexual Reproduction –** reproduction involving two parents – male and female.
- **Asexual Reproduction –** reproduction involving only one parent. Offspring genetically the same as the parent.
- **Cloning –** asexual reproduction.

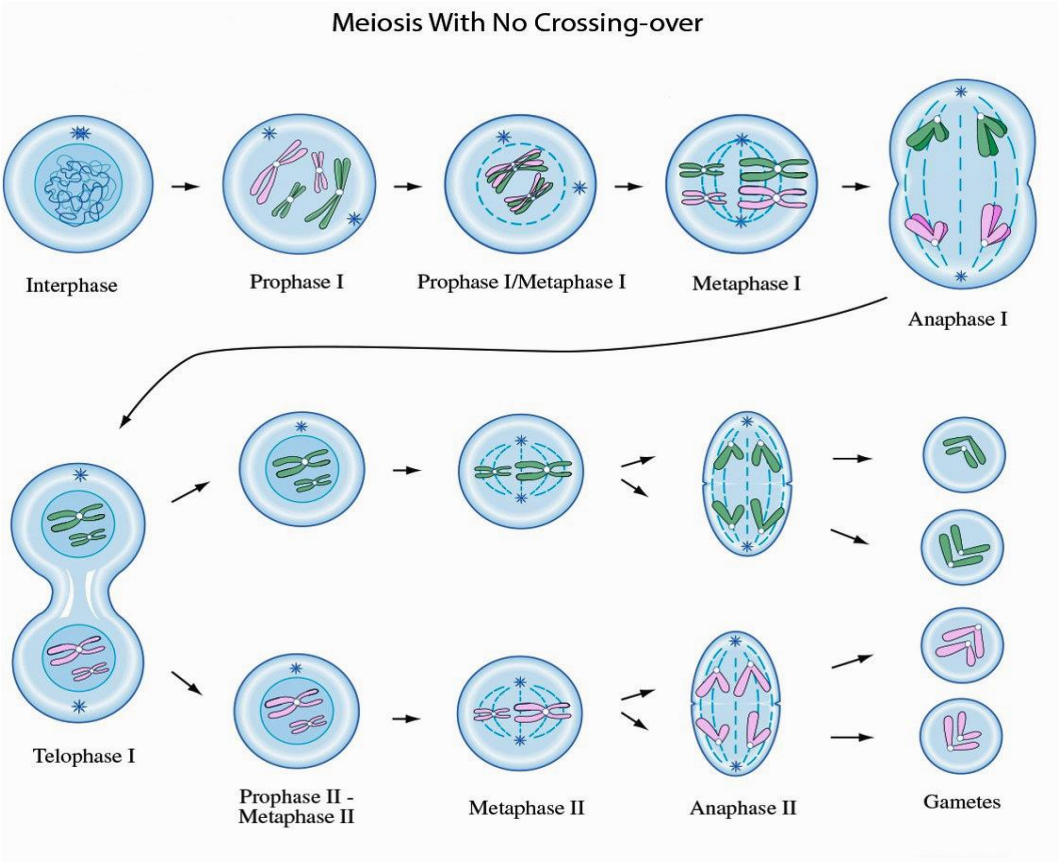
Stages of Meiosis I



- The first meiotic phase is **prophase 1**.
- As in mitosis, the nuclear membrane dissolves, chromosomes develop from the chromatin, and the centrosomes push apart, creating the spindle apparatus.
- The tight pairing of the homologous chromosomes is called **synapsis**
- These paired up chromosomes—two from each parent—are called **tetrads**.
- The point the points of contact, the physical link, between two (non-sister) chromatids belonging to homologous chromosomes is the **chiasmata**



- Homologous (similar) chromosomes from both parents pair up and **may** exchange DNA in a process known as **crossing over.** This results in genetic diversity.
- In **metaphase 1**, some of the spindle fibers attach to the chromosomes' centromeres.
- The fibers pull the tetrads into a vertical line along the center of the cell.
- **Anaphase 1** is when the tetrads are pulled apart from each other, with half the pairs going to one side of the cell and the other half going to the opposite side.
- It is important to understand that whole chromosomes are moving in this process, not chromatids, as is the case in mitosis.
- At some point between the end of anaphase 1 and the developments of **telophase 1**, cytokinesis begins splitting the cell into two daughter cells.
- In telophase 1, the spindle apparatus dissolves, and nuclear membranes develop around the chromosomes that are now found at opposite sides of the parent cell / new cells.

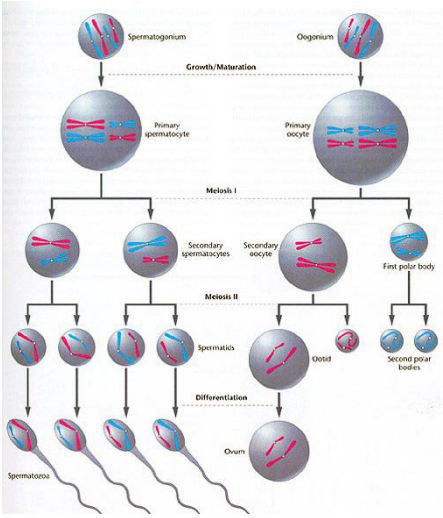


Stages of Meiosis II

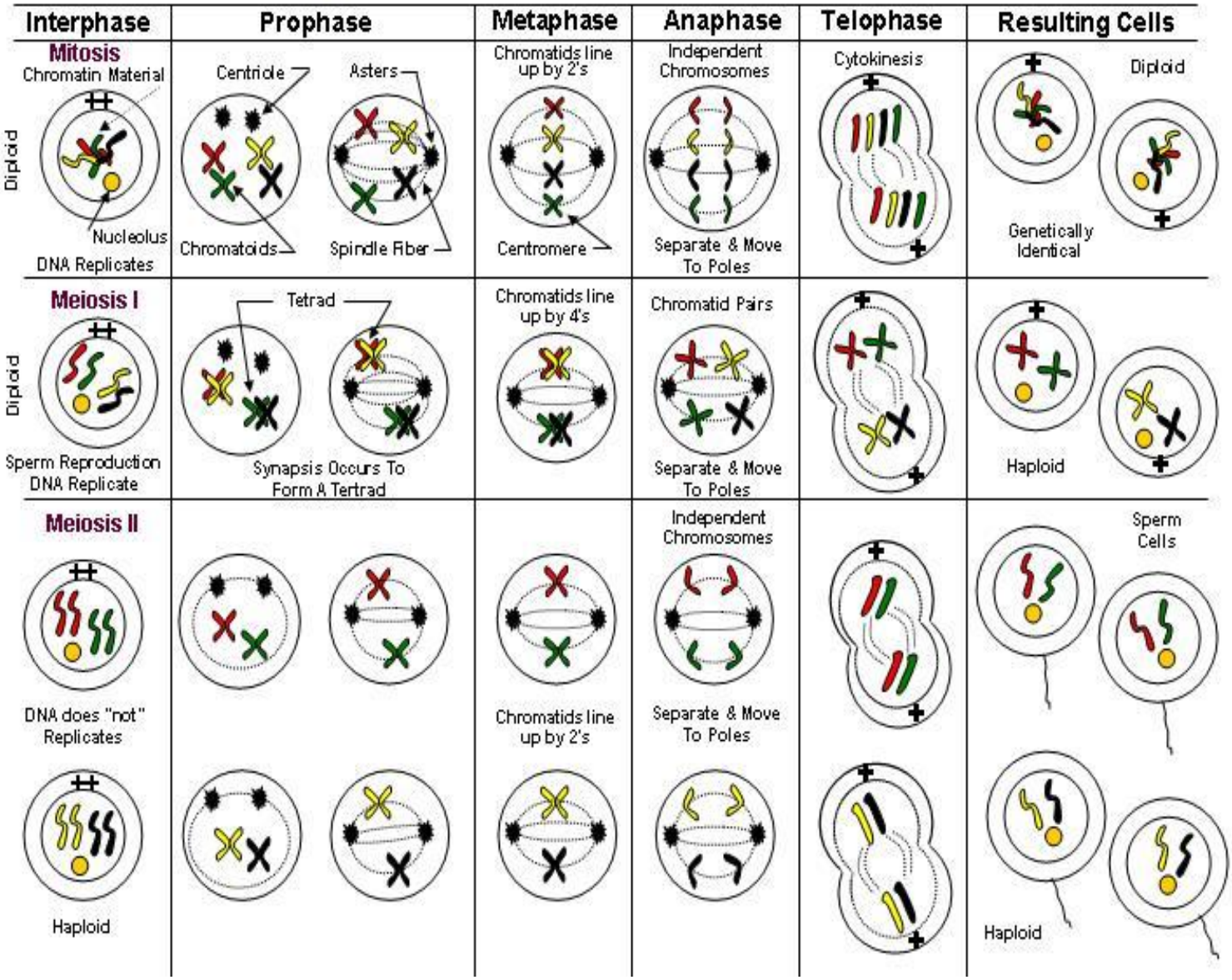
- In **prophase 2**, centrosomes form and push apart in the two new cells.
- A spindle apparatus develops, and the cells' nuclear membranes dissolve.
- Spindle fibers connect to chromosome centromeres in **metaphase 2** and line the chromosomes up along the cell equator.
- During **anaphase 2**, the chromosomes' centromeres break, and the spindle fibers pull the chromatids apart.
- The two split portions of the cells are officially known as "sister chromosomes" at this point.
- As in telophase 1, **telophase 2** is aided by cytokinesis, which splits both cells yet again, resulting in four haploid cells called gametes.
- Nuclear membranes develop in these cells, which again enter their own interphases.

Gamete Formation

- Spermatogenesis – 4 mature sperm cells from meiosis.
- Oogenesis – 1 egg and 3 polar bodies



Comparison of Mitosis and Meiosis



Note: no crossing over is shown in meiosis.

Mutations

- Mutation – any change in the DNA
- Gene mutation
- Chromosomal mutation
- Agents causing mutations – radiation, chemicals, excess heat

Genetic Disorders

- **Causes of mutations –** chemicals, radiation, temperature, viruses
- **Nondisjunction** – chromatids do not separate properly during meiosis. Individual formed from such gametes have extra or missing chromosomes. as Down's Syndrome
- **Trinucleotide repeats –** sequences of 3 nucleotides is repeated, often several times in a gene when too many repeats are formed – cause genetic disorders triplet nucleotides -repeated too often as Huntington's
- **Defective genes** – does not produce correct protein as sickle cell anemia (A & T traded places)
- **Genetic disorders and their causes** as nondisjunction (Down's syndrome), trinucleotide repeats (fragile X and Huntington's), defective genes (sickle cell anemia, hemophilia)
- **Human genetic disorders –** can be dominant, recessive, sex-linked, epistatic, variable expressed
- **Crossover frequency –** during meiosis, pieces trade places – determining frequency

Examples of Human Genetic Disorders

Autosomal Dominant

- **Huntington Disease** - degenerative brain disorder which results in loss of both mental and physical abilities-- adult onset generally
- **Marfan Syndrome** - disorder of connective tissue affecting the heart, blood vessels, lungs, eyes, bones, and ligaments
- **syndactyly** – webbing between toes and fingers
- **Polycystic Kidney Disease** - a disease which causes cysts to grow on a person's kidneys (and liver); the third leading cause of kidney failure in the United States
- **Brachydactyly** – short fingers
- **Myotonic Dystrophy –** a disorder that causes muscle weakness and the inability of muscles to relax after use.
- **porphyria** - a group of disorders caused by a deficiency of an enzyme in the pathway for making heme (a component of hemoglobin) -- this causes a variety of symptoms: sensitivity to light, mental changes which border on insanity, itchy and blistering skin, dark colored urine, abdominal pain and cramping, and hairiness
- **achondroplasia** – growth defect causing abnormal body proportions, the arms and legs are very short while the torso is normal in size.
- **chronic simple glaucoma** – increased pressure inside the eyeball
- **hypercholesterolemia** -excessive levels of cholesterol in the blood stream
- **polydactyly** – extra toes and fingers
- **Ehlers-Danlos Syndrome** - connective tissue disorders characterized by articular hypermobility (the ability to flex joints beyond the "normal" range), skin hyperelasticity (the ability to stretch the skin away from the body), and fragile skin and tissues (easy bruising and easily ruptured skin and blood vessels).
- **Neurofibromatosis** - trait characterized by cafe-au-lait ("coffee and milk" pigmented skin) spots and small tumor-like growths on or under the skin-- deformation of bones and curvature of the spine can also be symptoms

- **Nonsyndromic deafness –** hearing loss due to a defective gene. Most defects affect the structure of the inner ear
- **Congenital cataracts** – clouding of the lens in the eye
- **Familial high cholesterol** – high cholesterol levels

Autosomal Recessive

- **Tay-Sachs Disease** - a degenerative disorder causing death usually by age 5 – Jewish heritage
- **sickle cell anemia** - disease causing the red blood cells in the body to have a sickle shape (not a round shape). These sickle shapes can block veins, arteries, and capillaries and cause blood flow to an area to be stopped for a while. This can have serious side effects such as tissue death and stroke.
- **Beta thalassenia *(Cooley's Anemia)*-** a defect in the beta chain of hemoglobin resulting in severe anemia
- **galactosemia –** individuals lack the enzyme that helps the body break down galactose.
- **albinism *(Oculocutaneous)*** - disorder characterized by absence of pigment in hair, skin, and eyes
- **agammaglobulinemia –** defect that causes the absence of the white blood cells (B cells) causing recurrent bacterial infections
- **phenylketonuria** - individuals with PKU cannot digest the amino acid phenylalanine (part of many proteins)-- levels of phenylalanine rise in the bloodstream and cause brain damage
- **cystic fibrosis** - A disease caused by defective chloride transport that leads to high levels of mucus in the lungs and pancreas, high sweat chloride levels, and other digestive and respiratory problems.

Sex-linked

X- dominant

- **ichthyosis simplex** (d) Ichthyosis is a form of severe dry skin that causes affected areas to look like fish scales
- **Hypertrichosis -** generalized hairiness covering the whole body

X- linked recessive

- **hemophilia** (r) - There is a defect in blood coagulation factor VIII which prevents blood clotting. This causes hemorrhage, easy bruising, and prolonged bleeding from wounds.
- **red-green colorblindness** (r) – unable to distinguish between red and green
- **Duchenne's muscular dystrophy** (r) - a disease that begins to affect individuals between the ages 2 and 6. It causes muscle wasting and weakness. This can eventually affect all muscles of the body. Generally by age 10-12 affected individuals become confined to a wheelchair.
- **Anhidrotic Ectodermal Dysplasia -** a group of disorders characterized by the absence of sweat glands, abnormal teeth, and hypotrichosis (less hair than normal)
- **Fragile X Syndrome** - A disorder that causes various levels of mental impairment-- from learning disabilities to severe retardation, both combined with delayed speech and language development. caused by more than 200 repeats of the trinucleotide CGG. Karyotypes of individuals with Fragile X Syndrome appear to be missing a small piece of the X chromosome near the end.
- **Lesch-Nyhan Disease -** the absence of an enzyme HPRT (hypoxanthine-guanine phosphoribosyl transferase) causes an accumulation of uric acid in the urine and self-mutilative behavior

Y-linked

- **Hairy ears** - hair grows on the pinnae of the ears-- in some cases it is quite thick; in others it is only one
- or two long hairs

Nondisjunction

Autosomes:

- **Down's syndrome (trisomy 21**) - an extra copy of the 21st chromosome-- generally through
- non-disjunction but occasionally the extra copy of critical chromosome material can be as
- a result of translocation causing a combination of birth defects including some mental retardation
- and characteristic facial features
- **monosomy 21** – a chromosome missing in pair 21.

Sex Chromosomes:

- **monosomy X (Turner's syndrome)**
- **trisomy X**
- **XXY (Klinefelter's syndrome)**
- **XYY**
- **XXXX or XXXY**

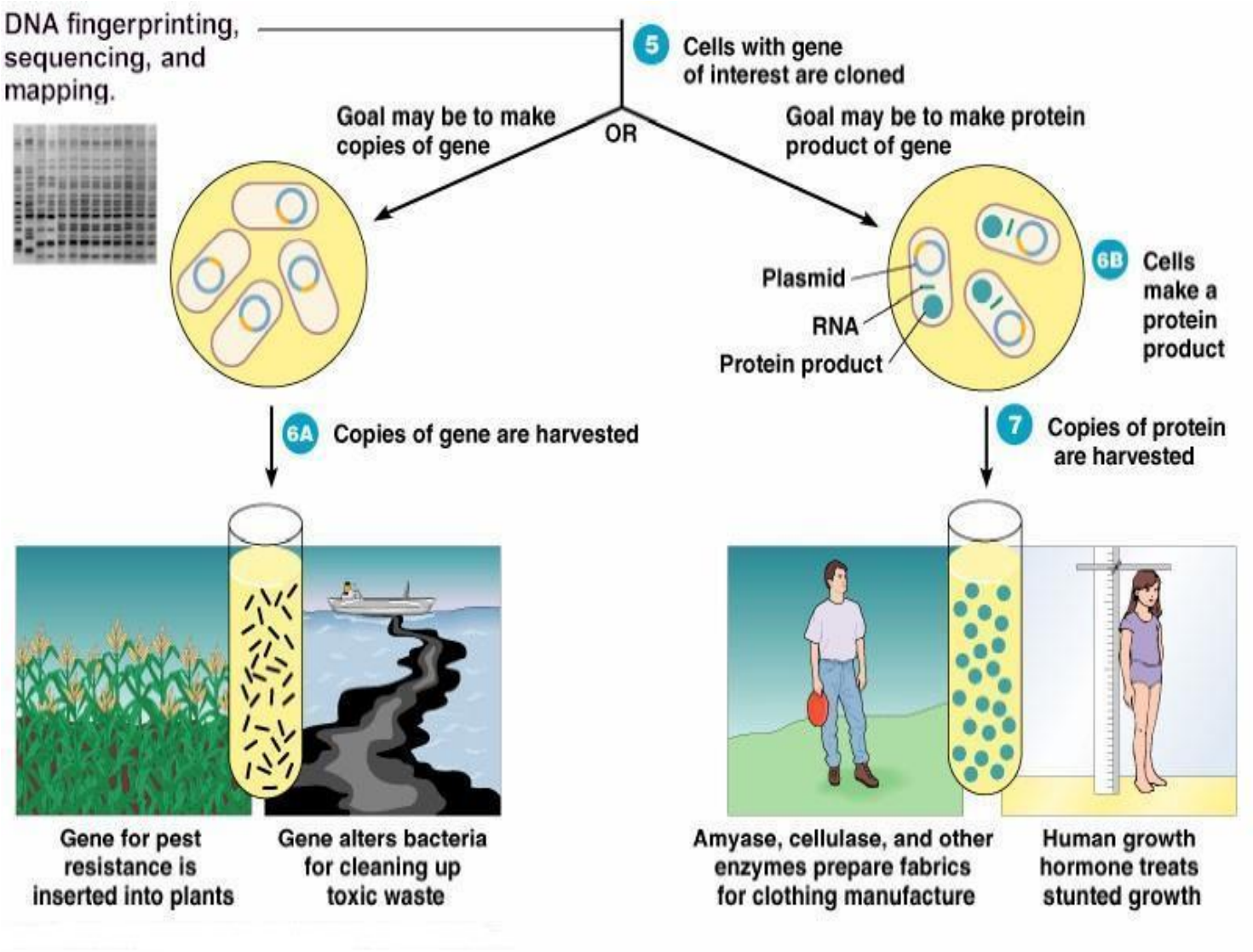
Multifactorial Inheritance (many genes + environment)

- **cleft palate and/or lip**
- **club foot**
- **congenital dislocation of the hip**
- **spina bifida (open spine)**
- **hydrocephalus (with spina bifida)**
- **pyloric stenosis**
- **diabetes mellitus** – Type I diabetes - Individuals do not produce insulin and must take shots of insulin to control their blood glucose levels
- **Breast Cancer** - multi-factorial- two genes have recently been identified inherited in an autosomal dominant fashion that increases susceptibility to breast cancer. These genes are known as BRCA1 and BRCA2. They are responsible for a very small proportion of all cases, particularly those which affect women at a younger age.

DESIGNER GENES - BIOTECHNOLOGY

- Technology to manipulate DNA – techniques often called **genetic engineering** or
- **Recombinant DNA Technology**-Technology used to manipulate DNA
- Procedures often called genetic engineering
- Recombinant DNA - DNA from two sources
- Transgenic individuals have DNA from another organism
- Often involve putting genes into viruses or bacteria.
- Vectors are the pieces of DNA used to transfer genes into a host cell – often plasmids of bacteria

Overview of Biotechnology



Basic Tools of DNA Technology

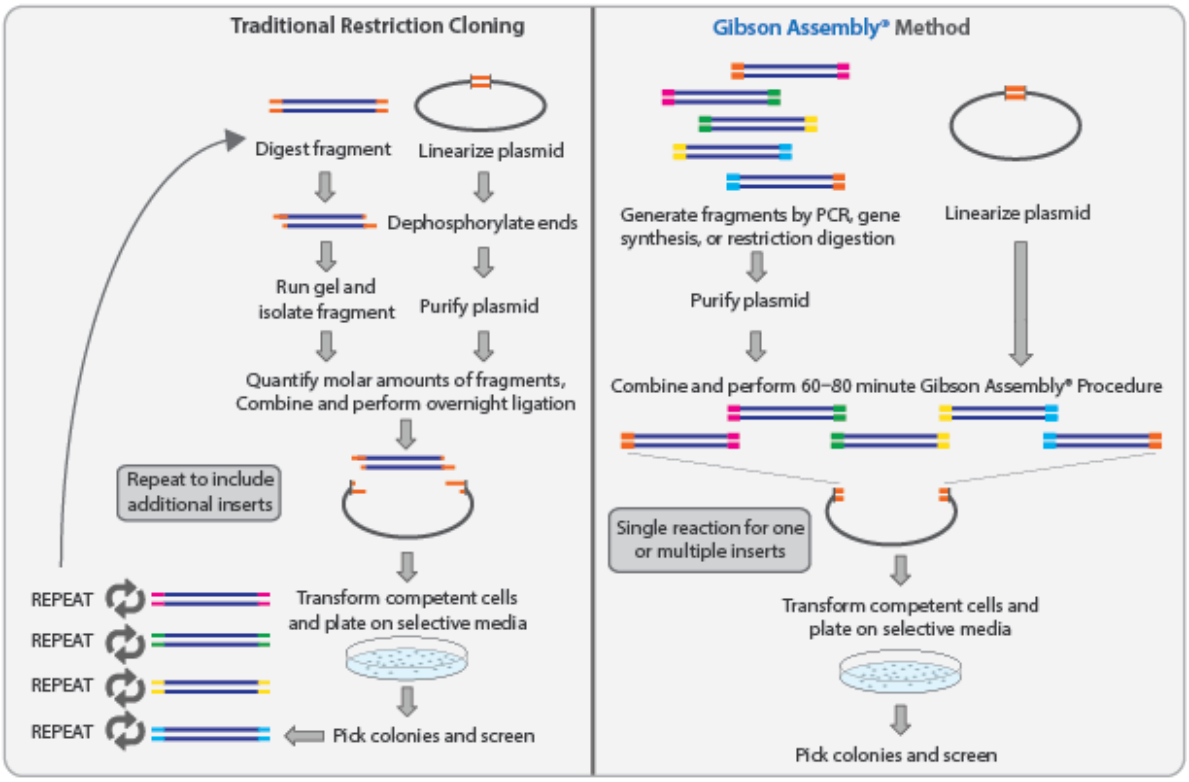
1. Identifying desired DNA
2. Cutting DNA with Restriction Enzymes
3. Inserting DNA into Vector as Plasmid
4. Connecting DNA pieces with Ligase
5. Inserting Vector into Host Cell as bacterium
6. Cloning desired DNA and Vectors
7. Storing clones in DNA Libraries
8. Identifying cloned genes with Radioactive Probes
9. Analyzing DNA by cutting fragments and separating by Electrophoresis

DNA Analysis Technologies

- **identifying –** recognizing desired DNA fragment or plasmid using **radioactive probes**
- **cutting DNA -** using desired **restriction enzymes** or " enzymatic scissors"
- **making hybrids of DNA** using **Hybridization techniques**
- **cloning DNA** – using other cells or in a test tube as with **PCR –** Polymerase Chain Reaction – clones - DNA segments in a test tube quickly and inexpensively. May use very small amounts of DNA to clone
- **storing DNA** in **DNA libraries** of plasmids or bacteriophages of genome DNA or cDNA.
- **separating DNA segments** with **electrophosesis**
- **transferring DNA** using **blotting**
- **imaging DNA** with **autoradiography**
- **analyzing DNA** by **sequencing** or determining the nucleotide sequence of a gene, **microassays** analyze gene function and expression, **DNA fingerprinting techniques** such as **RFLP** or restriction fragment length polymorphism, VNTRs or Variable Number Tandem Repeats, **STRs** or Short Tandem Repeats, **Ribosomal DNA Analysis**, or **Y-chromosome Analysis**

Cloning methods currently in use:

- **Traditional Restriction digestion cloning**
  - Plasmids and inserts are digested with the same restriction enzymes and then ligated together. See diagram below.
- **Gateway Recombination**
  - Regions of homology between insert DNA and plasmid are used for a recombination event that transfers the insert DNA into the plasmid.
- **Gibson assembly**
  - DNA fragments containing homologous overlapping ends are ligated together in one reaction.
- **TA cloning**
  - Linearized plasmids engineered to have single T overhangs are ligated together with a PCR product insert. Most DNA polymerases leave an A overhang on PCR products, which allows them to base pair with the T overhangs on the TA cloning plasmid**.**



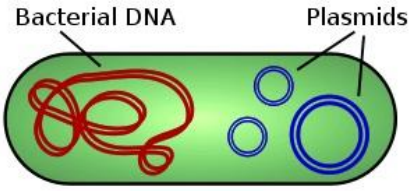
Basic Terminology

- **Recombinant DNA** – DNA from two different sources combined. Often involve putting genes
- into viruses or bacteria using a **vector**.
- **Inserting a gene into a bacterium** - Organism provides the desired piece of DNA which is spliced into a piece of DNA used to transfer the genes or **vector** which is inserted to a Host cell (often a bacterium)
- **Plasmids–** in bacteria, circular DNA serve as vectors. Easily taken up by bacterial cells. It is more difficult to insert vector into Eukaryotic cells.
- **Transgenic organisms** have DNA from another organism
- **Restriction enzymes -** enzymes to cut DNA at a particular spot and **DNA ligase** enzymes reattach ends.
- **Hybridization** – process of putting pieces of DNA together.
- **Chromosome mapping –** determining the location of genes on a chromosome and making a map of restriction sites as **Restriction Maps.**

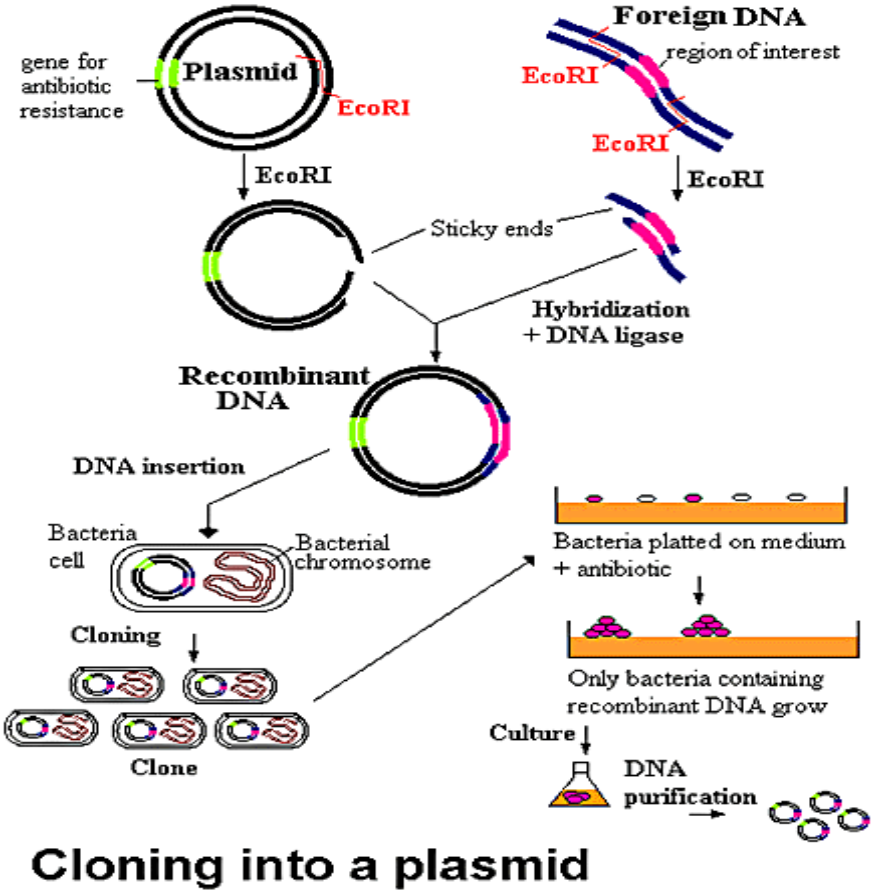
Basic Tools of Cloning

- **Gene selection** & isolation from Donor
  - Eukaryotic genes contain introns but bacteria do not contain the necessary enzymes to remove introns
  - Eukaryotic genes that are inserted into bacteria must be inserted without introns.
  - Use *reverse transcriptase* (from retroviruses) and modified M-RNA to produce **cDNA** with introns already removed

- **Plasmid selection & isolation**
  - A small DNA molecule that is physically separate from, and can replicate independently of, chromosomal DNA within a cell such as a bacterium
  - When used in genetic engineering – called **vectors**
  - Several methods to isolate plasmid DNA from bacteria



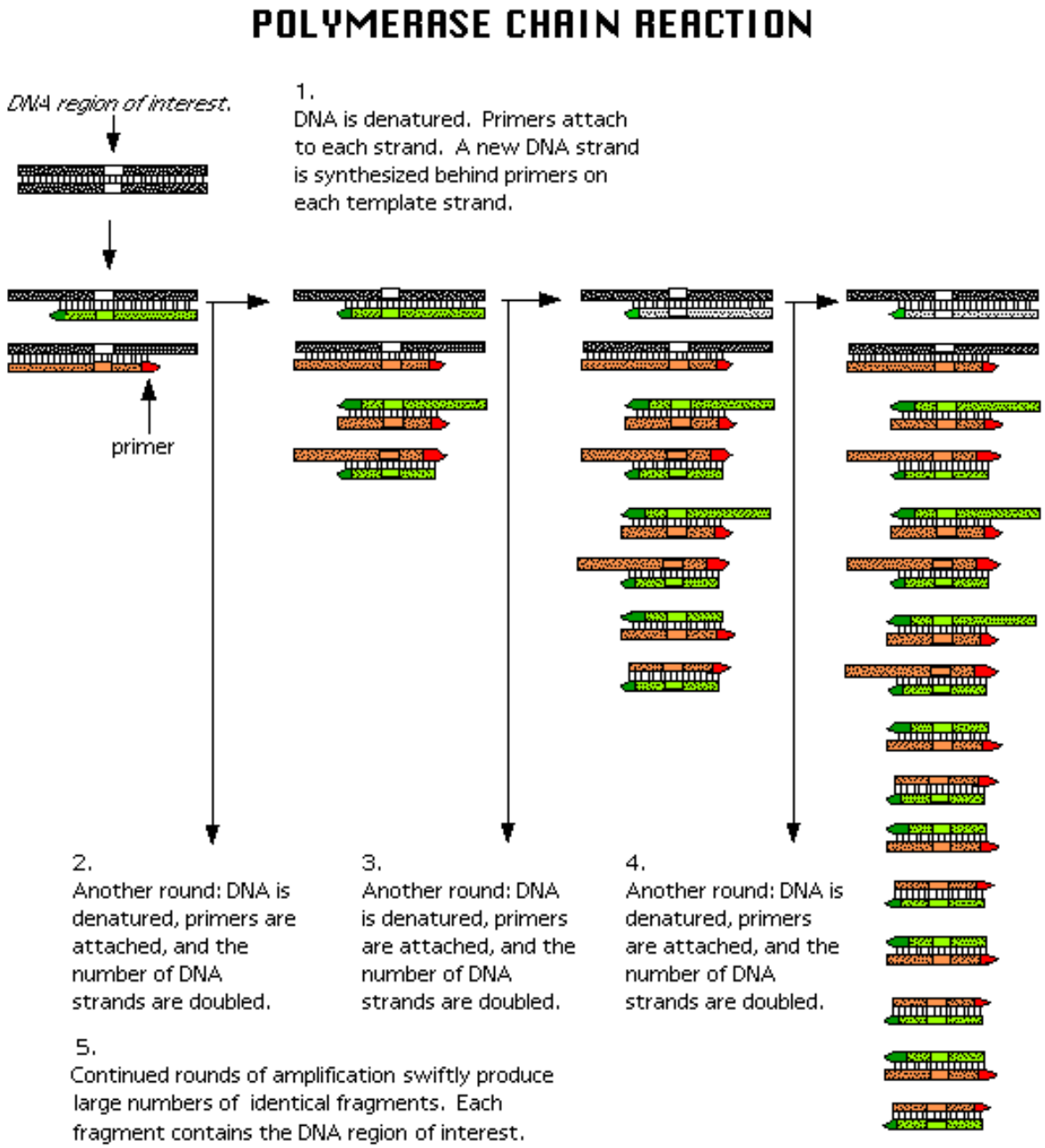
- **Restriction enzyme** to cut piece
- **Putting pieces together**
  - DNA hybridization
  - DNA ligase to reattach pieces
- **Insert** into Host bacteria
- **Clone** the bacteria



Cloning into a plasmid

Polymerase Chain Reaction (PCR)

- Technique for quickly making an unlimited number of copies of any piece of DNA
- Sometimes called "molecular photocopying



cDNA –

complementary DNA

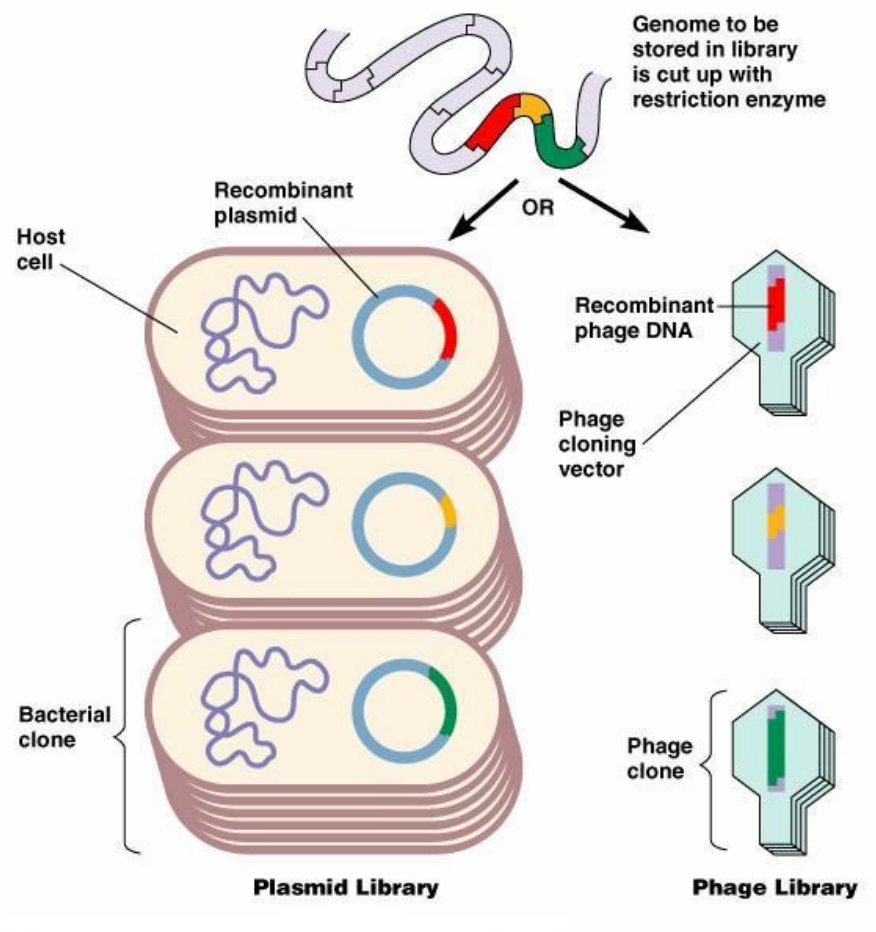
- cDNA- Eukaryotic genes contain introns but bacteria do not contain the necessary enzymes to remove introns
- Eukaryotic genes that are inserted into bacteria must be inserted without introns.
- Use *reverse transcriptase* (from retroviruses) and modified m-RNA to produce cDNA with introns already removed

Techniques for Storing, Identifying, Separating Clones

- Storing clones in DNA Libraries
- Cloning within cells and with PCR
- Identifying cloned genes with Radioactive Probes
- Analyzing DNA by cutting fragments and separating by Electrophoresis/nucleic acid hybridization/DNA probes
- Transferring DNA from gel by Blotting
- Imaging with autoradiography
- DNA Sequencing to determine exact sequence
- Microassays to analyze gene function

DNA Libraries

- **Genomic** – normal DNA
- **cDNA** – modified to remove introns
- Fragments are stored in plasmids or bacteriophages

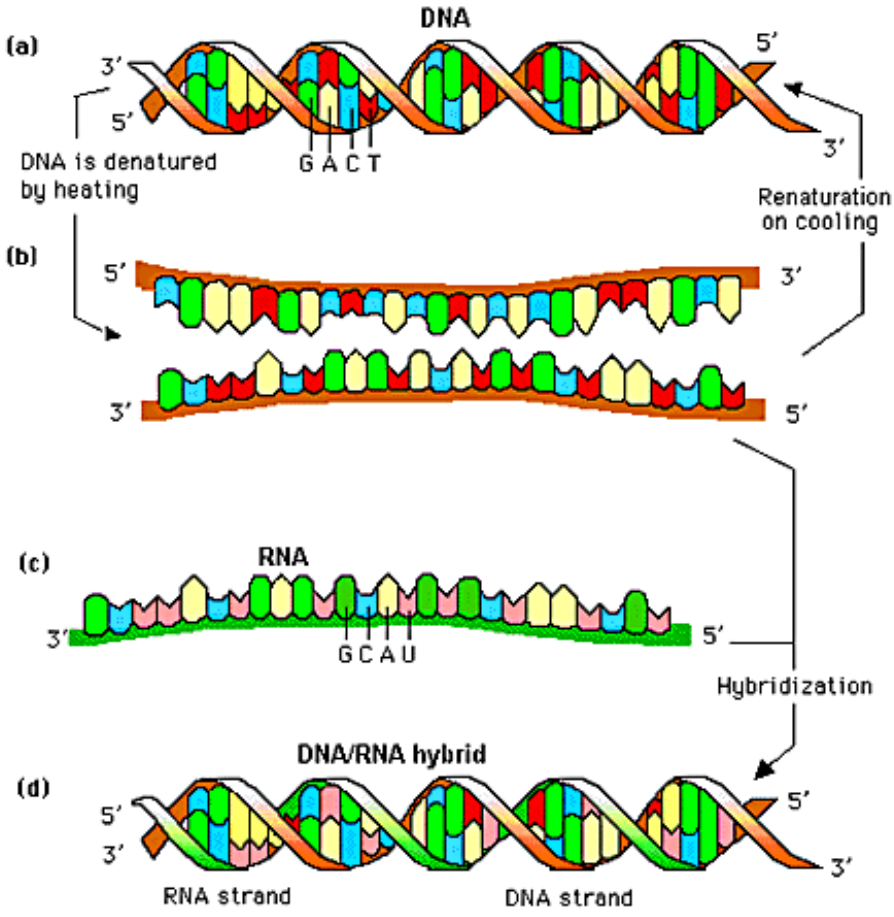


DNA Hybridization

- Base pairing of two single strands of DNA or RNA.
- Can be DNA-DNA, DNA-RNA
- Can be a radioactive probe

Radioactive Probes

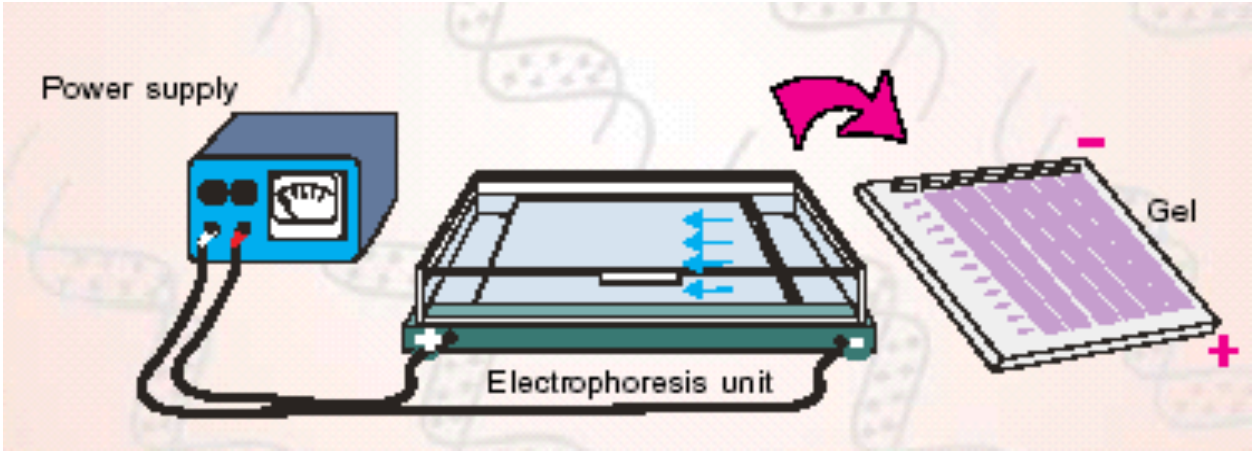
- short, radioactive strands of DNA
- will pair up with complementing strands of DNA
- fragments that contain the labeled pieces will show up on an x-ray film



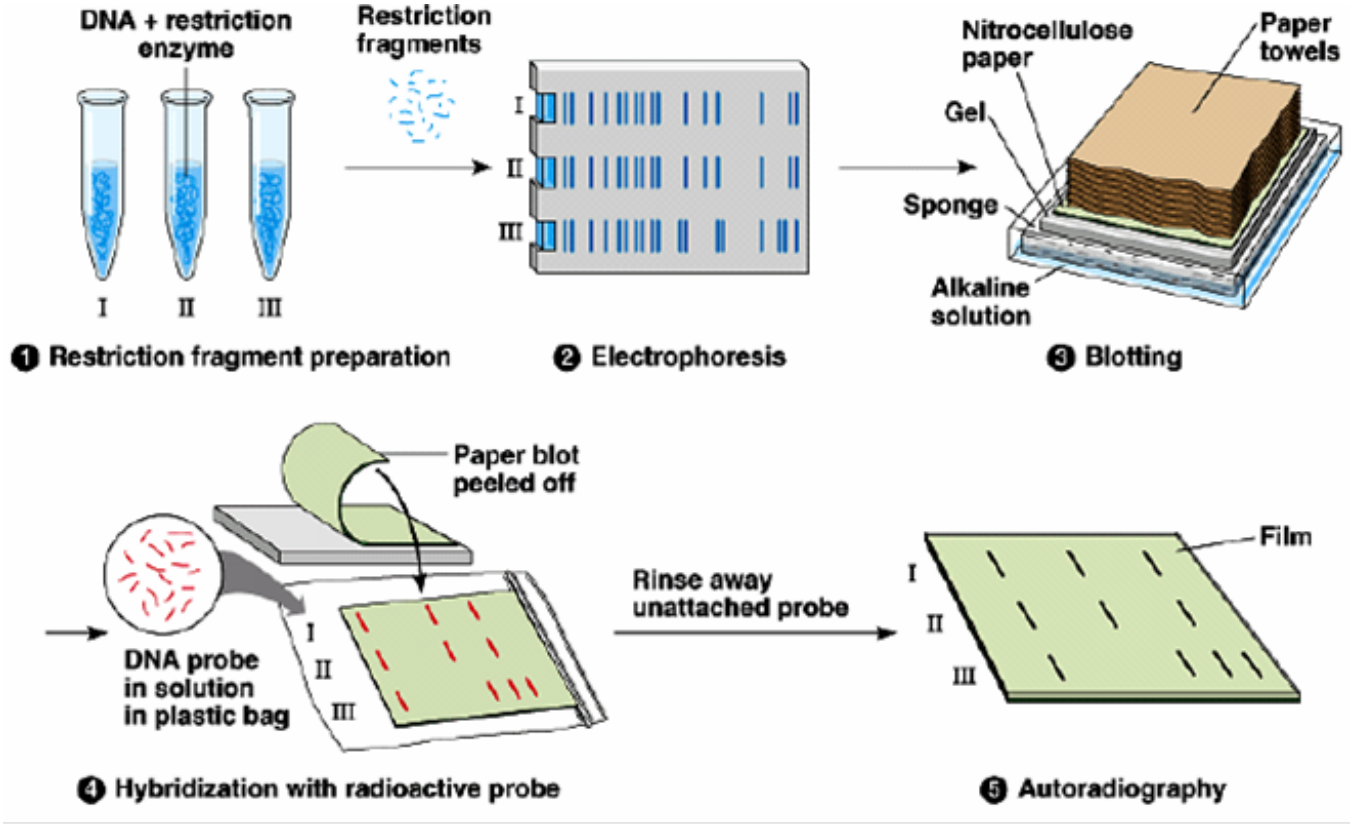
Nucleic Acid Hybridization

Electrophoresis

A process in which molecules (such as proteins, DNA, or RNA fragments) can be separated according to size and electrical charge by applying an electric current to them.



DNA Analysis

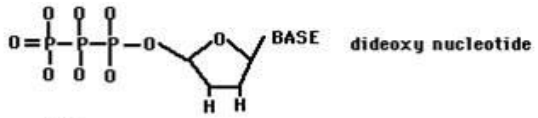


Analysis of DNA Fragments

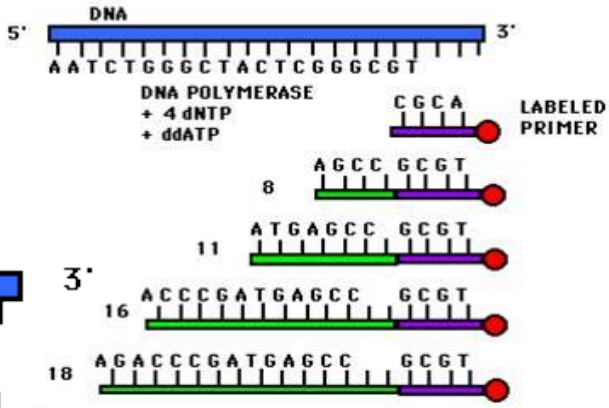
- When a plasmid is digested by restriction enzymes, the length of each fragment can be analyzed on a gel
- Then the physical map of the plasmid can be constructed.
- The DNA on a gel can be analyzed by hybridization after transfer onto a membrane, this is called a **Southern blot**.
- A similar procedure called a **Northern blot** is used to detect mRNA on a membrane.
- **Reverse transcription mediated PCR** can also be used to analyze mRNA from cells.

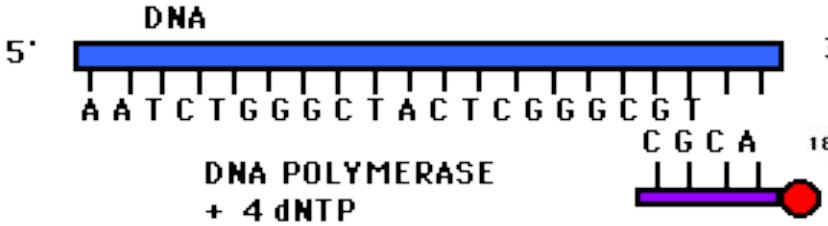
Sanger Method of DNA Sequencing

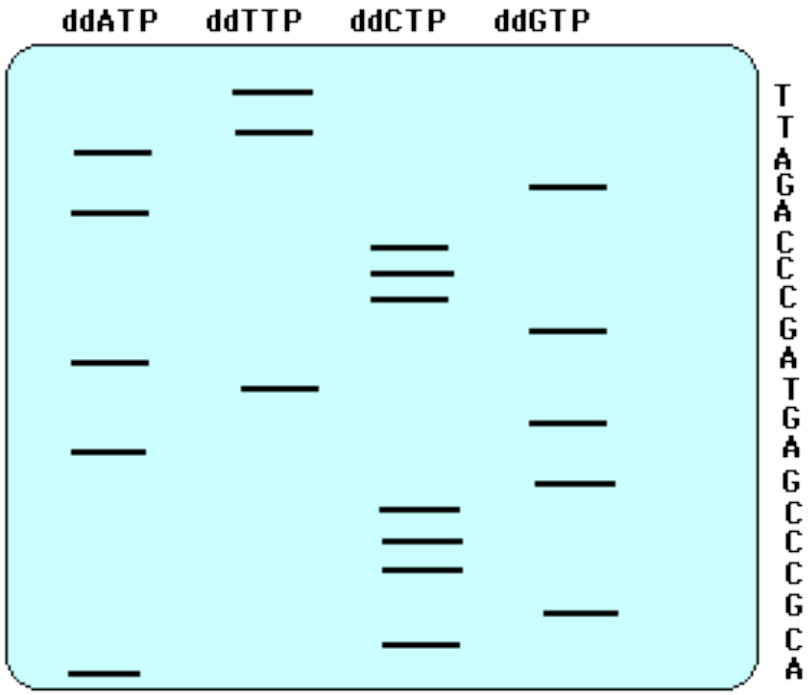
- Analysis of genes at the nucleotide level
- Tool has been applied to many areas of research
- Polymerase chain reaction (PCR) - quickly making an unlimited number of copies of any piece of DNA requires knowing the sequence of the piece to be copied
- Amino acid sequences can be determined more easily by sequencing a piece of cDNA
- Can utilize sequencing to identify the site of a point mutation
- Utilizes 2',3'-dideoxynucleotide triphospates (ddNTPs)



- First convert double stranded DNA into single stranded DNA
- Determine the **exact nucleotide sequence**
- Columns for A,T,C, and G
-
- The **sequence of the sequenced strand**, 5' to 3', is **AATCTGGGCTACTCGGGCGT**
- The sequence of the **strand of DNA complementary** to the sequenced strand is 5' to 3'







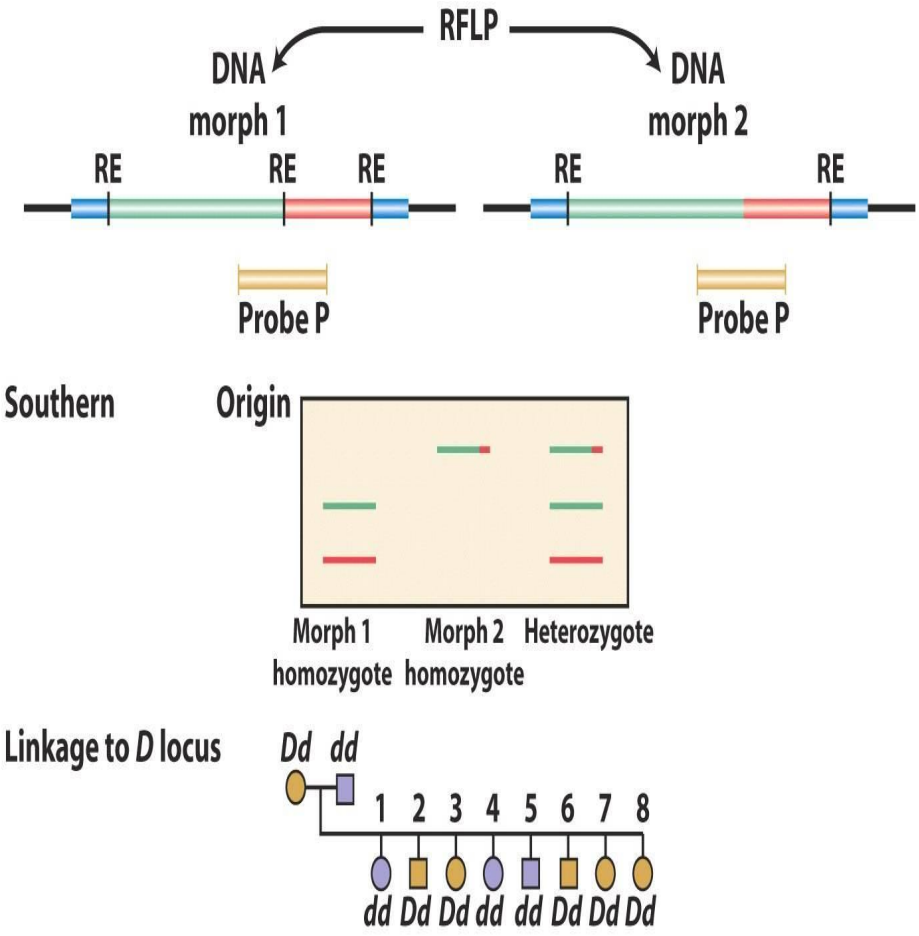
ACGCCCGAGTAGCCCAGATT

DNA Fingerprinting Techniques

- RFLP – Restriction Fragment Length Polymorphism (original)
- PCR – Polymerase Chain Reaction
- VNTRs – Variable Number Tandem Repeats
- STRs - Short Tandem Repeats
- Ribosomal DNA analysis
- Y-chromosome analysis

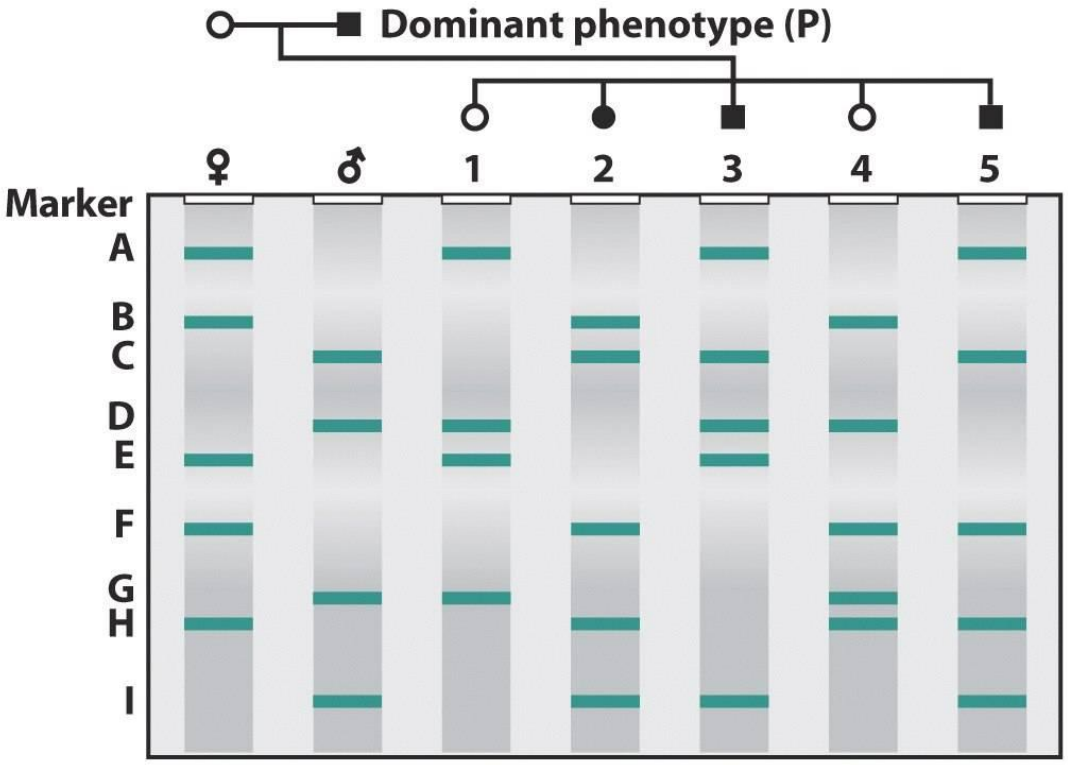
Restriction Fragment Length Polymorphism – RFLP

- The DNA of an organism is cut up into fragments using restriction enzymes.
- A large number of short fragments of DNA will be produced. (RFLP's)
- Electrophoresis is a technique used to separate the DNA fragments according to their size.
- Uses- identification of diseased genes including oncogenes, identification of viral infections, determining family relationships among individuals, and identifying tissue found at a crime scene.
- Genetic variations at the site where a restriction enzyme cuts a piece of DNA.
- Such variations affect the size of the resulting fragments.
- These sequences can be used as markers on physical maps and linkage maps.



VARIABLE NUMBER TANDAM REPEATS (VNTR's)

- Short nucleotide sequences
- Organized in clusters of tandem repeats
- VNTR = 14-100 base pairs
- SNR = 2- 10 base pairs

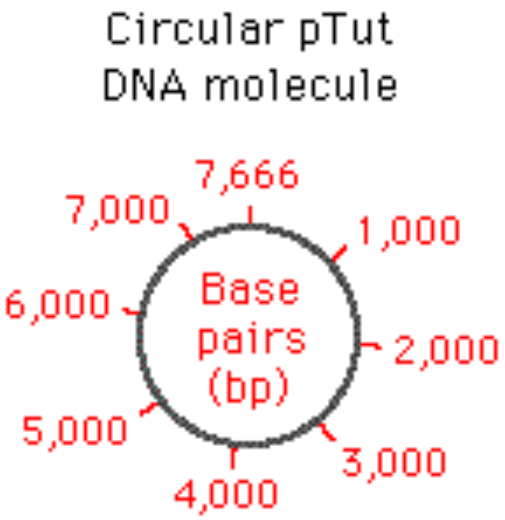


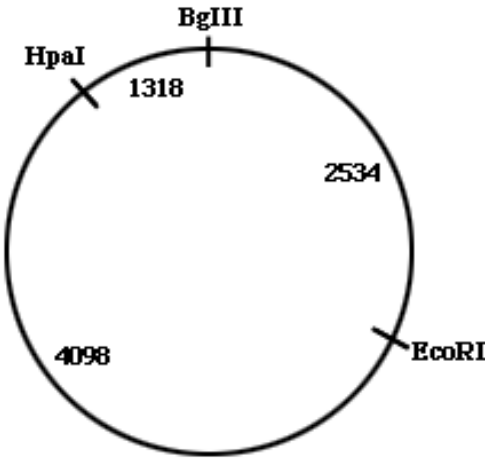
ANALYSIS EXAMPLES

F and H Always inherited together — linked?
A and B In progeny, always either A *or* B — "allelic"?
A and D Four combinations; A and D, A, D, or neither — unlinked?
F, H, and E Always *either* F and H *or* E — closely linked in trans?
Allele *P* Possibly linked to I and C.

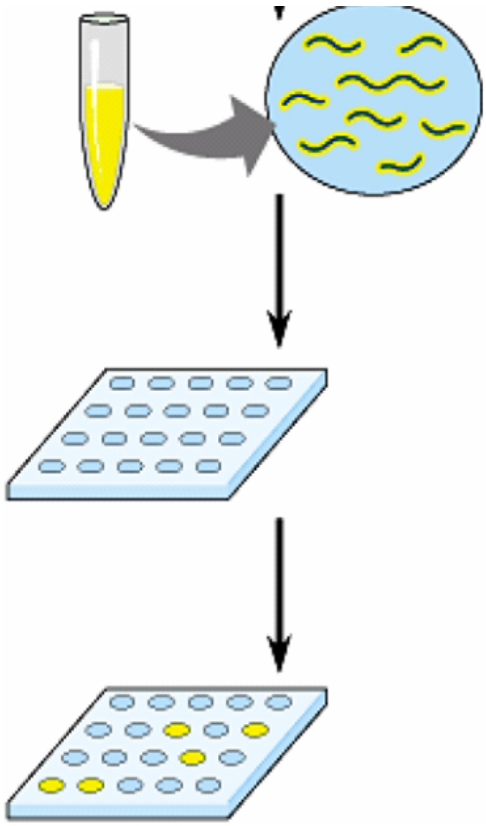
Restriction Mapping

- Description of **restriction enzyme cleavage sites** within a piece of DNA
- Use of **different restriction enzymes** to analyze and generate a physical map of genomes, genes, or other segments of DNA





DNA Microassay



- studying **how large numbers of genes interact with each other**
- precisely apply tiny droplets containing functional DNA to glass slides
- attach fluorescent labels to DNA from the cell they are studying.
- labeled probes are allowed to bind to complementary DNA strands on the slides
- slides are put into a scanning microscope that can measure the brightness of each fluorescent dot
- brightness reveals how much of a specific DNA fragment is present, an indicator of how active it is.
- **Advantages of using microarray technology**:
  - Readily available mature technology
  - Standard methods available

o Relatively inexpensive

- **Limitations of microarray technology**:
  - Dynamic range of measurement is limited:
    - Intensity of fluorescent dyes
    - Sensitivity of scanning instruments
    - Non-specific hybridization

o Known genome.

RNA-Seq

- RNA-seq refers to the method of using Next Generation Sequencing (NGS) technology to sequence RNA which can then be used to quantify relative levels of transcription for every gene in an organism's genome across different conditions
- NGS technology is an ultra-high-throughput technology to measure DNA sequences.
- **Advantages of RNA-seq over microarray** include:
  - Wider measurable range of expression levels
  - Not dependent on known genome
  - Free of hybridization artifacts
  - Can be used to identify and quantify small regulatory RNAs
  - Possibility of one platform for all applications

Next Generation Sequencing Platforms

- Roche 454 sequencer
- Illumina Genome Analyzer (Solexa sequencing)
- Applied Biosystems SOLiD sequencer
- Ion Proton sequencing
- PacBio real-time sequencing
- Comparison of the Second-generation DNA sequencing technologies

Platform	Chemistry	Read Length	Run Time	Gb/Run	Advantage	Disadvantage
454 GS Junior (Roche)	Pyro-sequencing	500	8 hrs.	0.04	Long Read Length	High error rate in homopolymer
454 GS FLX+ (Roche)	Pyro-sequencing	700	23 hrs.	0.7	Long Read Length	High error rate in homopolymer
HiSeq (Illumina)	Reversible Terminator	2*100	2 days (rapid mode)	120 (rapid mode)	High-throughput / cost	Short reads Long run time (normal mode)
SOLiD (Life)	Ligation	85	8 days	150	Low Error Rate	Short reads Long run time
Ion Proton (Life)	Proton Detection	200	2 hrs.	100	Short Run times	New*
PacBio RS	Real-time Sequencing	3000 (up to 15,000)	20 min	3	No PCR Longest Read Length	High Error Rate

Applications of Biotechnology Techniques

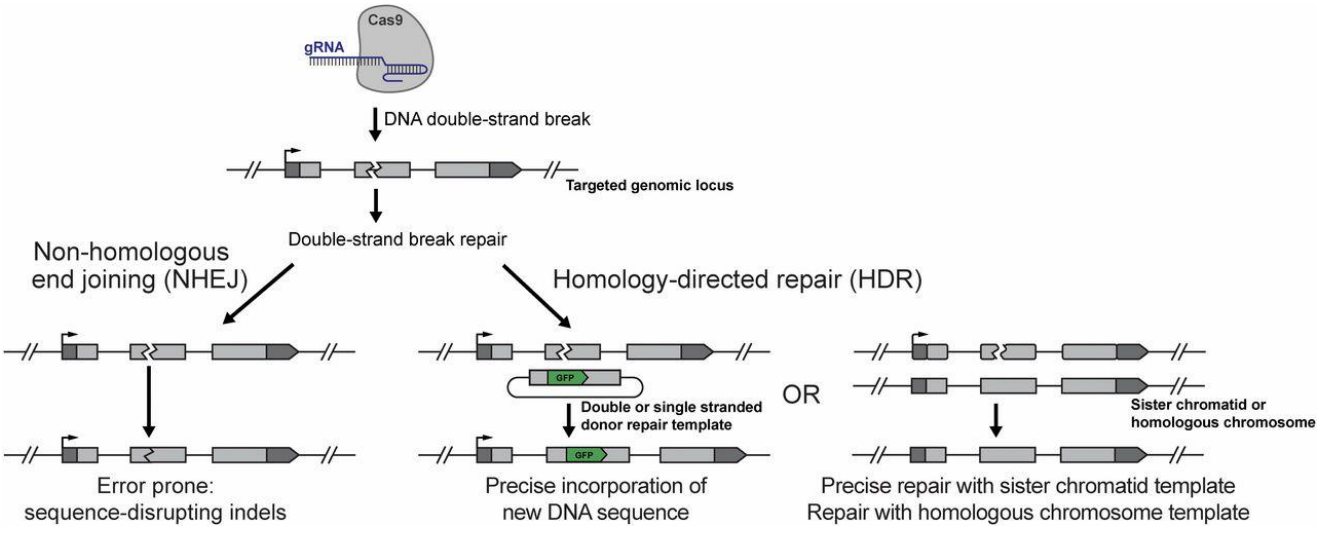
- Human Genome Project - entire gene make up of humans
- Diagnosis of Disease – PCR & DNA probes
- Human Gene Therapy
- Vaccines & Pharmaceutical Products
- Forensics – DNA Fingerprints (RFLP & VNTR)
- Environmental – Recycling & detoxification
- Agricultural – transgenic organisms

Gene Therapy - changing the expression of a person's genes - body (somatic) or germ cells done *in vitro* or *ex vitro*.

CRISPR Cas9 gene editing

CRISPR are Clustered Regularly-Interspaced Palindromic Repeats.

- They are comprised of repeated DNA sequences separated by "spacers."
- Originally discovered in bacteria, they became more interesting to microbiologists when it was discovered that the spacer DNA between the repeats are derived from foreign DNA (like that of a bacteriophage) that has previously infected the bacterium or one if its ancestors.
- Having the foreign DNA inserted into the bacterial genome allows for the production of anti-sense RNAs that can be used to guide attacks on future infections with the same phage DNA.
- This has been compared to the function antibodies play in the adaptive immune system of higher organisms.



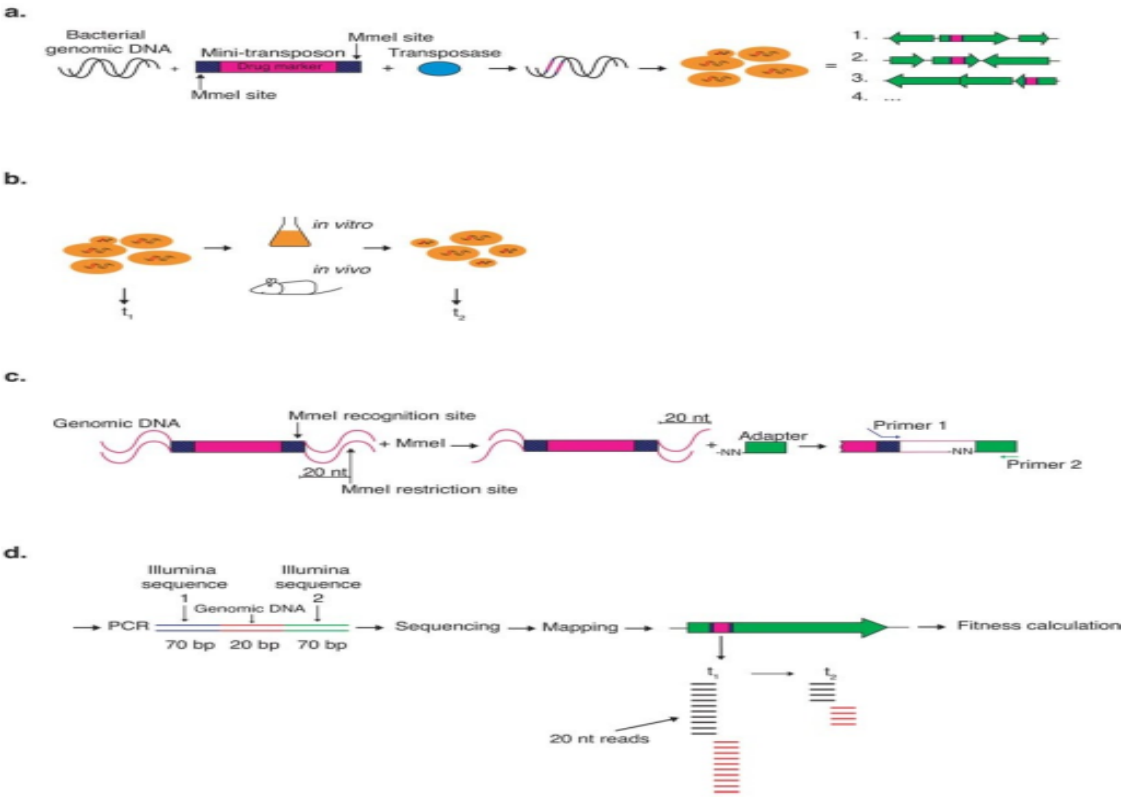
- Integrating these foreign DNAs into the bacterial genome to generate CRISPR spacers requires specialized enzymes that are coded in genes near the CRISPR called CRISPR associated (Cas) genes.
- Cas9 is an example of one of these genes that is being extensively used to edit DNA in the lab.
- Cas9 is an endonuclease that cuts DNA at a very specific site near a 20bp sequence that is complimentary to a guide RNA present within the Cas9 protein.
- The guide RNA can be altered by an experimenter to very specifically target a cut site in any DNA sequence.
- When Cas9 cuts, it forms a double strand break that can be repaired by non-homologous end joining which results in a small deletion of DNA near the cut site.

- Alternatively when homologous DNA is present, homology directed repair occurs.
- This homologous DNA could be provided by a sister chromosome; however, an experimenter could also provide the DNA that will be used to repair the break.
- If this DNA contains small changes, then the changes will be included in the repair, which is why this technology holds much promise for genetic engineering.

TN-Seq

TN-Seq is a popular modern technology for determining genes that contribute to an organism's fitness in any particular environment. The following letters refer to the figure below.

a. A very large pool of mutants is made, each with a single transposon inserted randomly into its genome. The transposon disrupts the gene into which it inserts, generally rendering it non-functional.
b. The pool of mutants is exposed to some sort of selective conditions like growth in a particular medium or animal host. Samples of the mutant pool are collected before exposure to the condition (t_1) and after exposure (t_2)
c. Chromosomal DNA is purified from the samples of mutants and prepared for rapid sequencing using one of the available next generation sequencing technologies such that the DNA adjacent to the transposon is sequenced.
d. Sequencing reads are mapped to the genome of the organism that was mutated. Comparing the maps of transposon mutants before and after exposure to the condition can yield information about which genes contribute to the organism's ability to survive the condition. These genes are called fitness determinants.



The scheme summarizes what TN-Seq results can mean:

- Decreased numbers of transposon mutants in the output (t_2) sample means that the genes that were disrupted are important for surviving the condition.
- Having the transposon disrupt those genes made the organism less fit, so they are underrepresented in the pool of mutants after exposure to the condition.

- Increased numbers of transposon mutants in the output (t_2) sample means that the genes that were disrupted are not important for surviving the condition.
- Having the transposon disrupt those genes made the organism more fit so they are overrepresented in the pool of mutants after exposure to the condition.

- No changes in numbers of transposon mutants in the output (t_2) sample means that the genes that were disrupted are not involved in surviving the condition.
- Having the transposon disrupt those genes didn't make the organism any more or less fit so they are unchanged in their representation in the pool of mutants after exposure to the condition.

Bioethics

Major concerns concerning safety and ethics of recombinant DNA technology.

- Potential Hazards vs. Potential Gains
- Concerns:
    - genetically modified foods
    - genetically engineering microbes
    - cloning whole organisms
    - embryonic stem cell research
    - gene therapy
    - genetic testing
    - bioterrorism

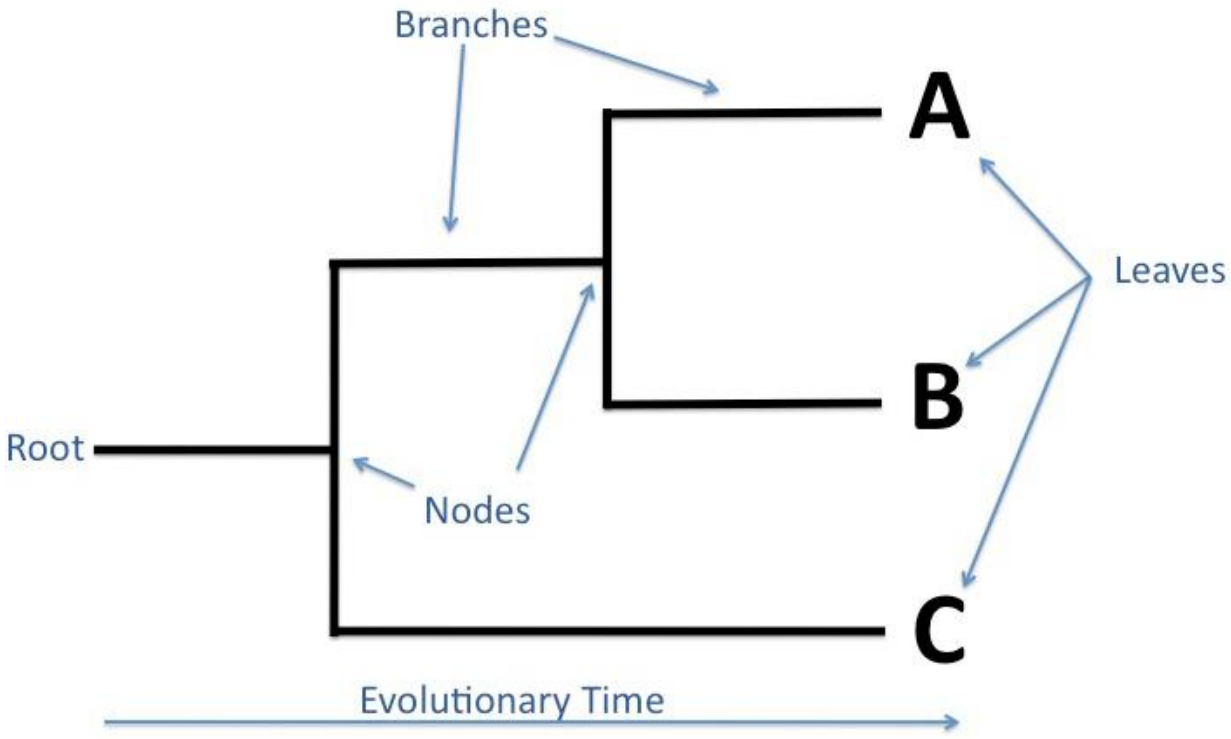
Epigenetics

The study of heritable changes in gene activity that occur without a change in the sequence of the genetic material. Epigenetics literally means 'in addition to genetics'.

- Epigenetic factors can regulate the amount of gene activity, influencing the growth and appearance of an organism
- There are several epigenetic ways in which gene activity can be prevented or controlled, including
    - modification of histone proteins
    - DNA methylation
    - RNA interference
- For any of these methods of gene regulation, the absence of the protein product of the gene causes a change in the function or development of the cell
- malfunctions in epigenetic control of gene activity have been implicated in cancer, cardiovascular disease and several inherited genetic conditions

Phylogenetics

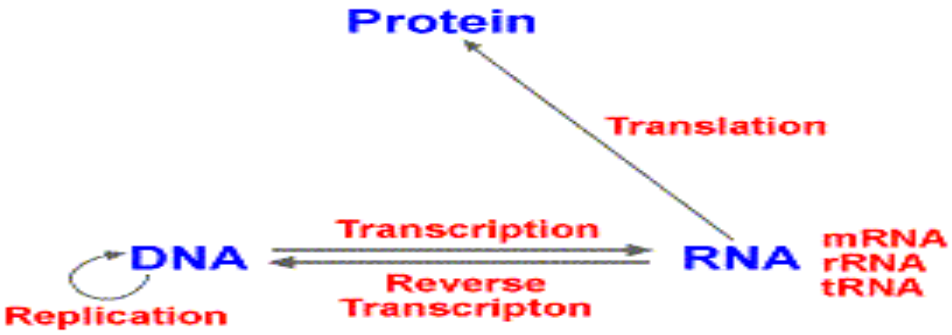
- Study of evolutionary relationships among groups of organisms based upon their genetics
- Has taxonomy folks in a turmoil – they can't agree so we have national lists for our taxonomy events



- Relationships between organisms are often represented using phylogenetic trees.
- Organisms more closely related to each other are physically closer on the tree.
- Common ancestors are represented by nodes that connect branches together.

DESIGNER GENES-MOLECULAR GENETICS

CENTRAL DOGMA OF MOLECULAR GENETICS



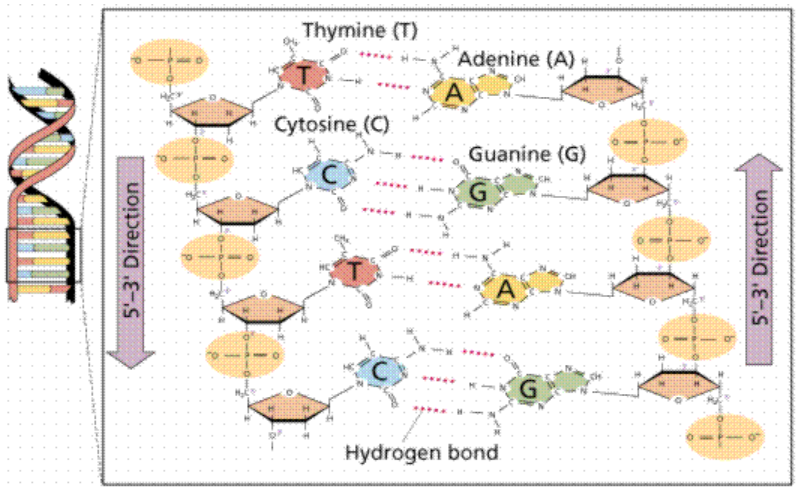
DNA ----□ RNA ---□ PROTEIN SYNTHESIS

REPLICATION□ TRANSCRIPTION□ TRANSLATION

- **Central dogma of molecular genetics is DNA -□ RNA -□ Protein.**
- **Exceptions among viruses – RNA to DNA (retroviruses) -** Exception is in retroviruses where genetic storage vehicle is RNA. It then makes a DNA which replicates to form double stranded DNA and continues through dogma.

DNA Structure

- **DNA structure** – double helix with **sugar** (deoxyribose), **phosphate** and **nitrogen bases** (**Adenine**, **Thymine**, **Guanine**, and **Cytosine**).
- Pairing is A with T and G with C

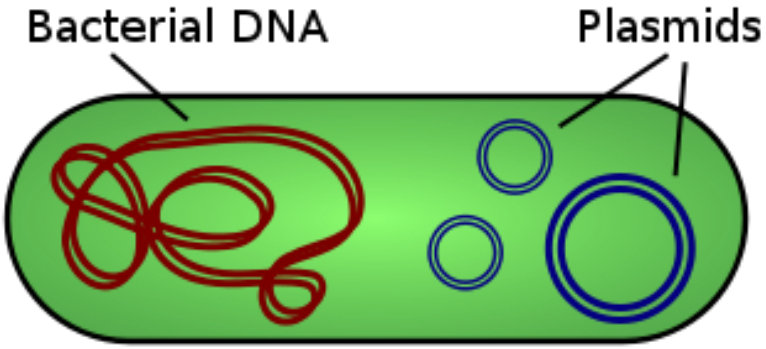


Nucleotide - basic unit of sugar, phosphate and nitrogen base - 4 kinds of nucleotides because of the 4 types of bases

Similarities between Prokaryotic and Eukaryotic DNA Replication

- Both prokaryotic and eukaryotic DNA replications occur before entering the nuclear division.
- Both prokaryotic and eukaryotic DNA replication works upon double-stranded DNA.
- The unwinding of both prokaryotic and eukaryotic DNA is done by DNA helicase.
- The unwound DNA strands are stabilized by single-stranded DNA-binding proteins (SSB).
- Both prokaryotic and eukaryotic DNA replication are multistep processes, which are carried out by an enzyme complex called DNA polymerases.
- Each type of DNA polymerases works in the 5' to 3' direction.
- RNA primers are required for the initiation of both types of DNA replications.
- The synthesis of the RNA primer is done by the enzyme called primase.
- Both prokaryotic and eukaryotic DNA replications occur in a semi-conservative manner where one old strand of DNA and one new strand of DNA can be found in the daughter cell.
- Both prokaryotic and eukaryotic DNA replications are bi-directional since the replications progress in both ways.
- Leading and lagging strands are formed in both types of DNA replications.
- The lagging strand produces the small DNA fragments called Okazaki fragments, which are eventually joined together.
- The time taken for both types of replications are around one hour.

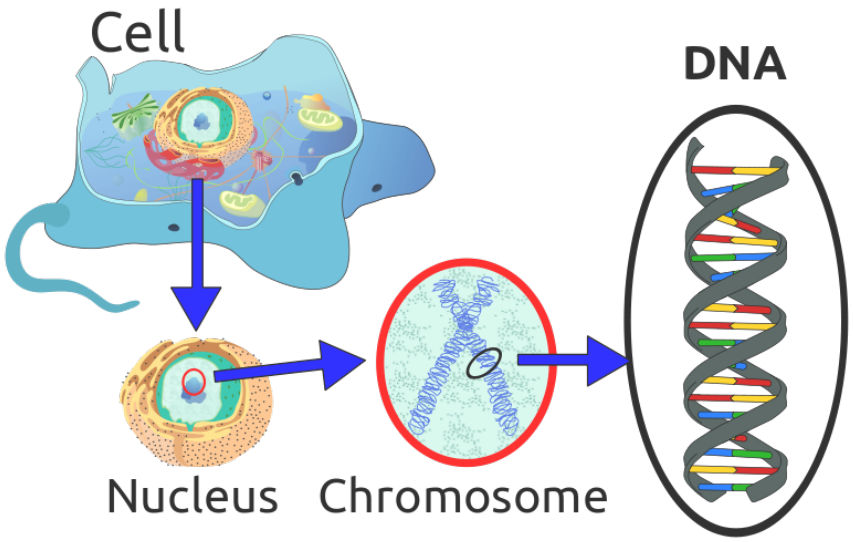
Prokaryotic DNA



Features of Plasmids

- Plasmids can be readily isolated from bacterial cells.
- They are self-replicative inside cells.
- They are composed of unique restriction sites for one or more restriction enzymes.
- The insertion of a foreign DNA fragment may not alter the replication properties of plasmids.
- Plasmids can be sequentially transformed into different types of cells and the transformants can be selected based on the antibiotic resistance properties of the transformed plasmids
- Plasmids can be used as vectors that carry foreign DNA molecules into both eukaryotic and prokaryotic cells.

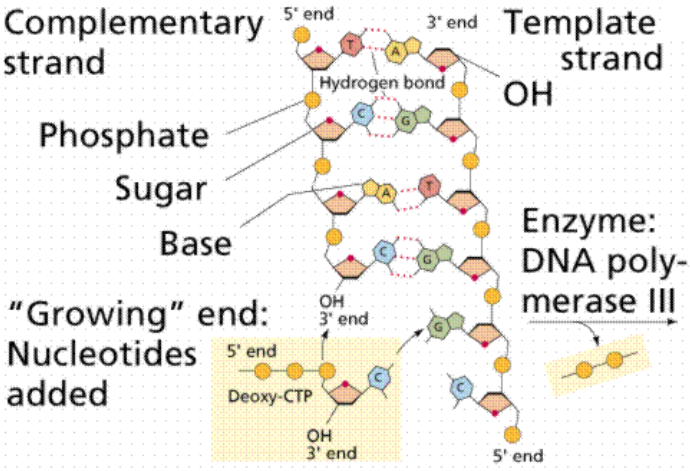
Eukaryotic DNA



PROKARYOTIC vs.	EUKARYOTIC
Structure-Closed Circular double stranded DNA	**Structure**-Linear double-stranded DNA with 2 ends
In the central portion of the cytoplasm	Inside the nucleus (some in mitochondria and some in chloroplasts)
Do not form the typical chromosome	Form chromosome or chromatin in the nucleus
Do not interact with histone proteins	Associated with histone proteins
Forms loop-like structures by wrapping around histone-like protein molecules but no nucleosome formation	Form distinct structural repeats called nucleosomes and shows higher order packaging
Quantity of DNA is small	Quantity of DNA is 50 times more than Prokaryotic
Only very few proteins interact with prokaryotic DNA	Large numbers of proteins interact with eukaryotic DNA
Usually codes for 300 to 500 proteins	Codes for thousands of proteins
Majority of DNA is coding, only very little non-coding region occurs in prokaryotic DNA	Majority of DNA is non-coding. Size of coding region is less than the non-coding regions.
Replication takes place in cytoplasm	**Replication** takes place in nucleus
Continuous process	Occurs during S phase of the cell cycle
Single origin of replication	Multiple origins of replication (over 1000)
Carried out by DNA polymerase I and II	Carried out by DNA polymerase α, δ, and ε.
The Okazaki fragments are comparatively large, 1000-2000 nucleotides in length.	The Okazaki fragments are small, around 100-200 nucleotides in length
DNA gyrase is required	DNA gyrase is not required
Rapid process – around 2000 nucleotides are added per second	Slow process – around 100 nucleotides added per second
Final product is 2 circular chromosomes	Final product is 2 sister chromatids
Introns absent in the coding region of DNA	Introns occur in the coding region of DNA
Transcription produces-mRNA unit is polycistronic – codes two or more proteins	**Transcription** produces- mRNA unit is monocistronic – codes only one protein

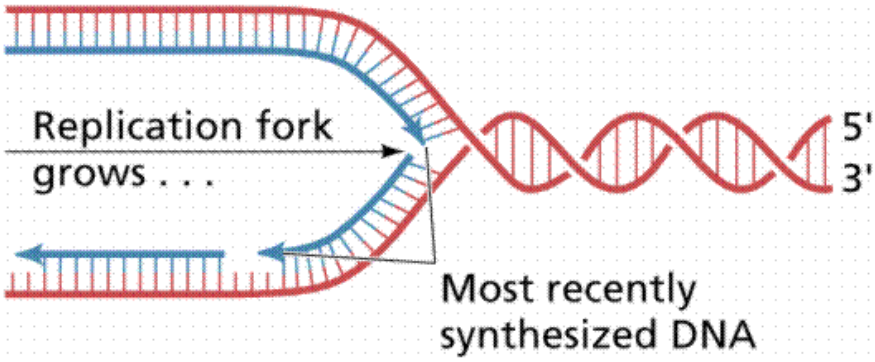
DNA Replication

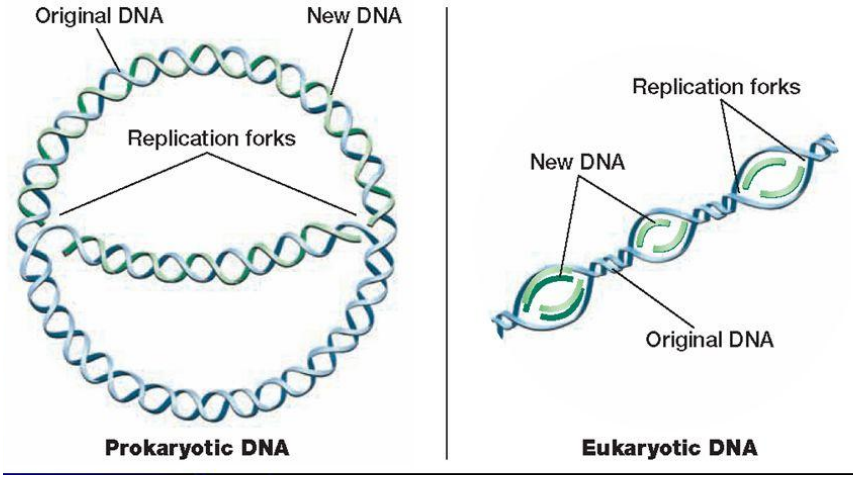
DNA replication is semi-conservative.



Events that occur:

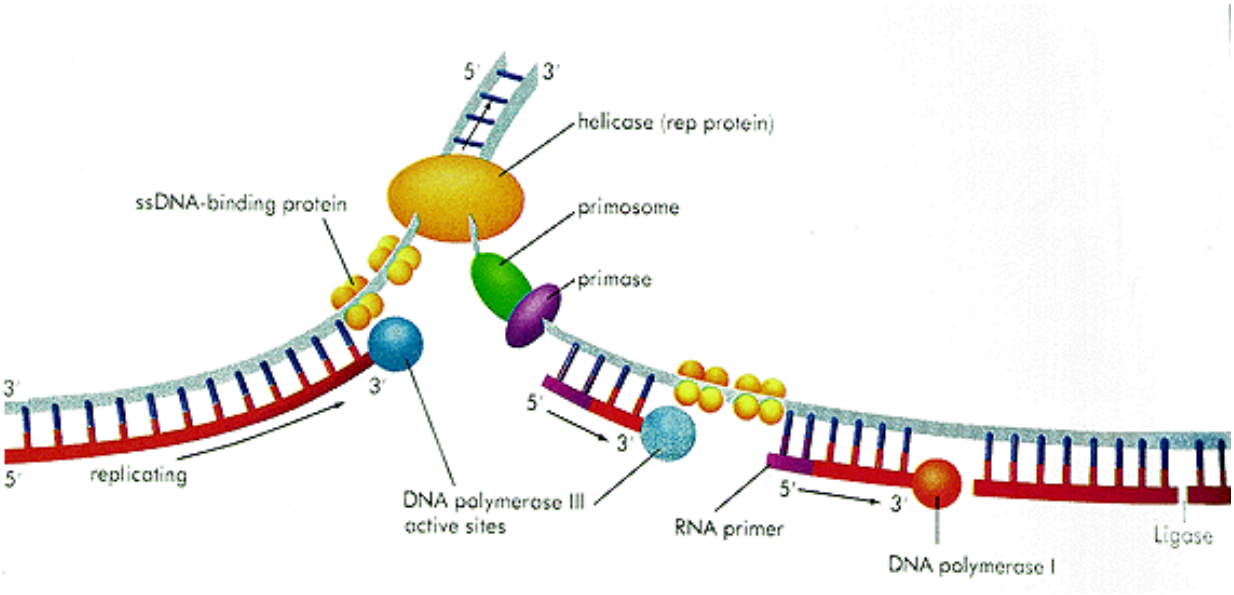
- **DNA polymerase** is the key enzyme
- DNA uncoils and splits
- template strand is read 3' to 5'
- new complementary strand must add new nucleotides to the 3' end – leading strand (continuous) while lagging strand is fragments (**Okazaki fragments**) latter attached with the enzyme ligase





Prokaryotic DNA Replication

- Prokaryotic DNA replication is the process by which prokaryotes such as bacteria and archaea duplicate their genome into a second copy, which can be transformed into a daughter cell.
- Prokaryotes consist of a double-stranded circular DNA molecule in their cytoplasm.
- Prokaryotic DNA comprises a single origin of replication. DNA helicase unwinds the DNA at the origin of replication by breaking the hydrogen bonds between the nitrogenous bases. The resultant Y-shaped structure is called the **replication fork.**
- Since prokaryotic DNA contains a single origin of replication, only two replication forks are formed during the replication process. These two replication forks process bi-directionally.
- The single-strand DNA-binding proteins (SSB) stabilizes the two unwound strands, which serve as the template strands for the replication.
- The enzyme, RNA primase synthesizes a five to ten base pairs long RNA primer, which is complementary to the template strand.



- Three types of DNA polymerases are involved in the prokaryotic DNA replication; DNA polymerase I, II, and III.
- Both initiation and elongation of the prokaryotic DNA replication are carried out by DNA polymerase III.
- The DNA polymerase III adds nucleotides in 5' to 3' direction.
- Due to the antiparallel nature of the DNA double-helix, one strand runs from 5' to 3' direction (leading strand). The other strand runs from 3' to 5' direction (lagging strand).
- Since, the lagging strand requires RNA primers continuously in order to synthesize DNA in the 5' to 3' direction, new fragments of DNA called Okazaki fragments are continuously formed.
- The gap filling and DNA repair are carried out by DNA polymerase I and II.
- The RNA primer is removed by the DNA polymerase I.

ENZYMES INVOLVED IN REPLICATION – Prokaryotic

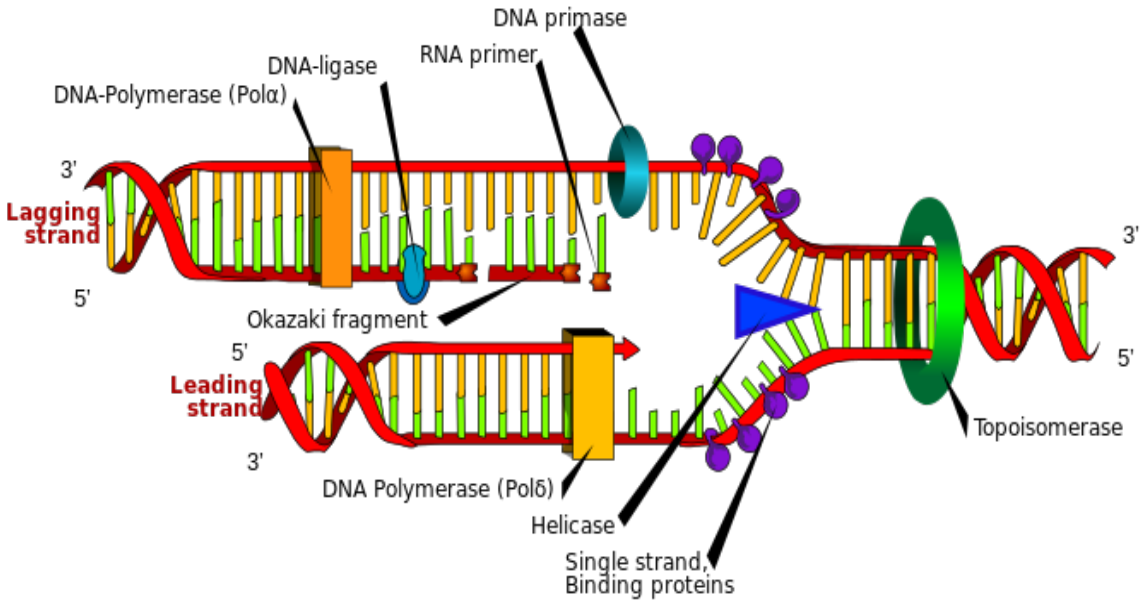
The **replication fork** is the unwound helix, with each strand being synthesized into a new double helix

- ***Topoisomerase*** is responsible for initiation of the unwinding of the DNA.
- ***Helicase*** accomplishes unwinding of the original double strand, once supercoiling has been eliminated by the topoisomerase.
- ***DNA polymerase (III)*** proceeds along a single-stranded molecule of DNA, recruiting free
- **dNTP's** (deoxy-nucleotide-triphosphates) to hydrogen bond with their appropriate complementary dNTP on the single strand (A with T and G with C), and to form a covalent **phosphodiester bond** with the previous nucleotide of the same strand.

  DNA polymerases **cannot** start synthesizing **de novo** on a bare single strand. It needs a primer with a **3'OH** group onto which it can attach a dNTP DNA polymerase also has proofreading activities, so that it can make sure that it inserted the right base, and nuclease (excision of nucleotides) activities so that it can cut away any mistakes it might have made.

- ***Primase*** attaches a small RNA primer to the single-stranded DNA to act as a substitute **3'OH** for DNA polymerase to begin synthesizing from. This RNA primer is eventually removed and the gap is filled in by **DNA polymerase (I).**
- ***Ligase*** can catalyze the formation of a phosphodiester bond given an unattached but adjacent 3'OH and 5'phosphate. This can fill in the unattached gap left when the RNA primer is removed and filled in.
- ***Single-stranded binding proteins*** are important to maintain the stability of the replication fork. Single-stranded DNA is very labile, or unstable, so these proteins bind to it while it remains single stranded and keep it from being degraded.

Eukaryotic DNA Replication



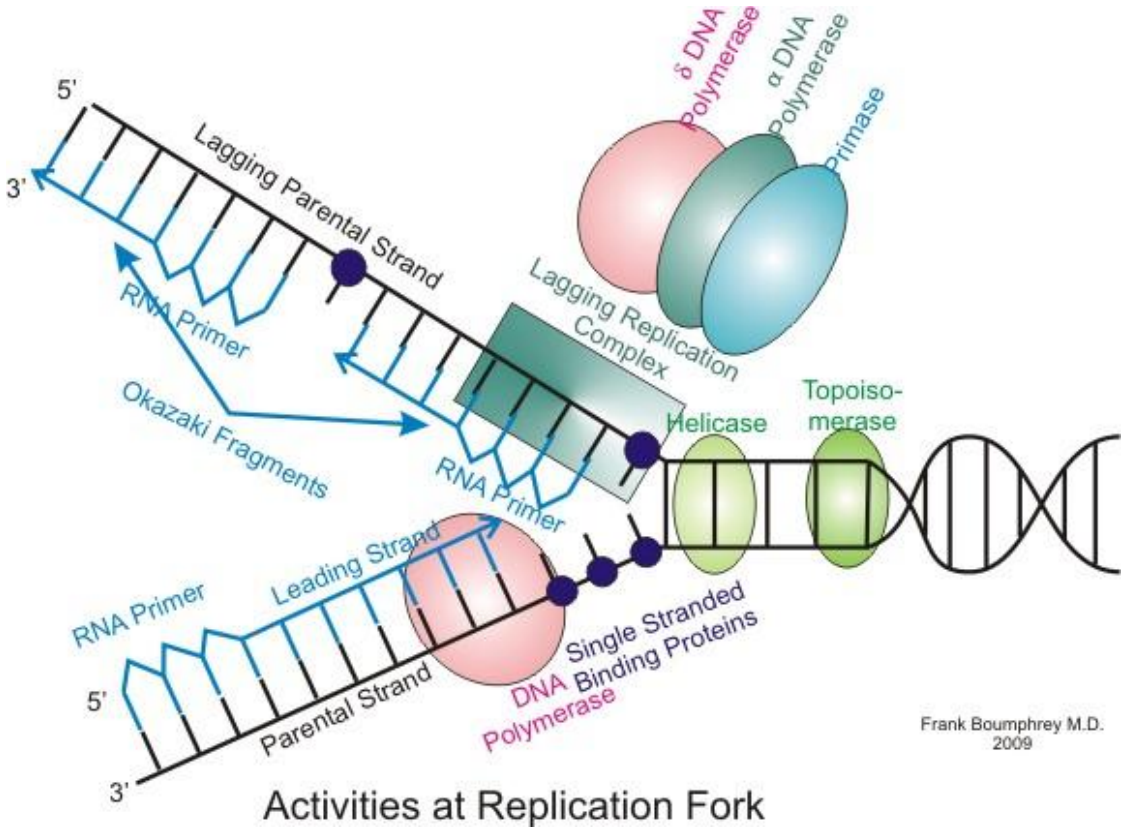
- Eukaryotic DNA replication is the process by which the eukaryotic genome duplicates prior to cell division. Though the basic mechanism of the eukaryotic DNA replication is similar to prokaryotic DNA replication, there are some differences due to the size and the structure of eukaryotic DNA.
- Eukaryotic DNA is double-stranded linear molecules.
- The amount of the eukaryotic DNA is around 50 times more than the prokaryotic DNA.
- Moreover, eukaryotic DNA is tightly packed with histones inside the nucleus of the cell.
- Therefore, DNA replication occurs in three steps; initiation, elongation, and termination.

Initiation

- The eukaryotic DNA replication occurs through multiple replication origins.
- The multiple replication origins form several replication bubbles per chromosome.
- DNA helicase and SSBs are involved in the unwinding and stabilization of the two templates at each origin of replication.

Elongation

- During elongation, DNA polymerases add new nucleotides to the existing 3' ends.
- The three types of DNA polymerases which are involved in eukaryotic DNA replication are DNA polymerase α, δ, and ε.
- The DNA polymerase α initiates the DNA replication whereas the DNA polymerase δ and ε are involved in the elongation.
- DNA polymerase α also requires an RNA primer to synthesize the new DNA strand and the primer is removed by the DNA polymerase β.
- The leading and lagging strands are formed in the same manner as in prokaryotic DNA replication. Eukaryotic DNA replication elongation is shown in *figure 2.*



Activities at Replication Fork

Termination

- Once the leading strand of a one replication bubble meets a lagging strand of a second replication bubble, the replication process is halted.
- Then, the RNA primer is removed, and the gap is filled by the freely-floating DNA polymerases. The nicks are joined by the DNA ligase.
- The multiple replication bubbles are shown in *figure 3*.

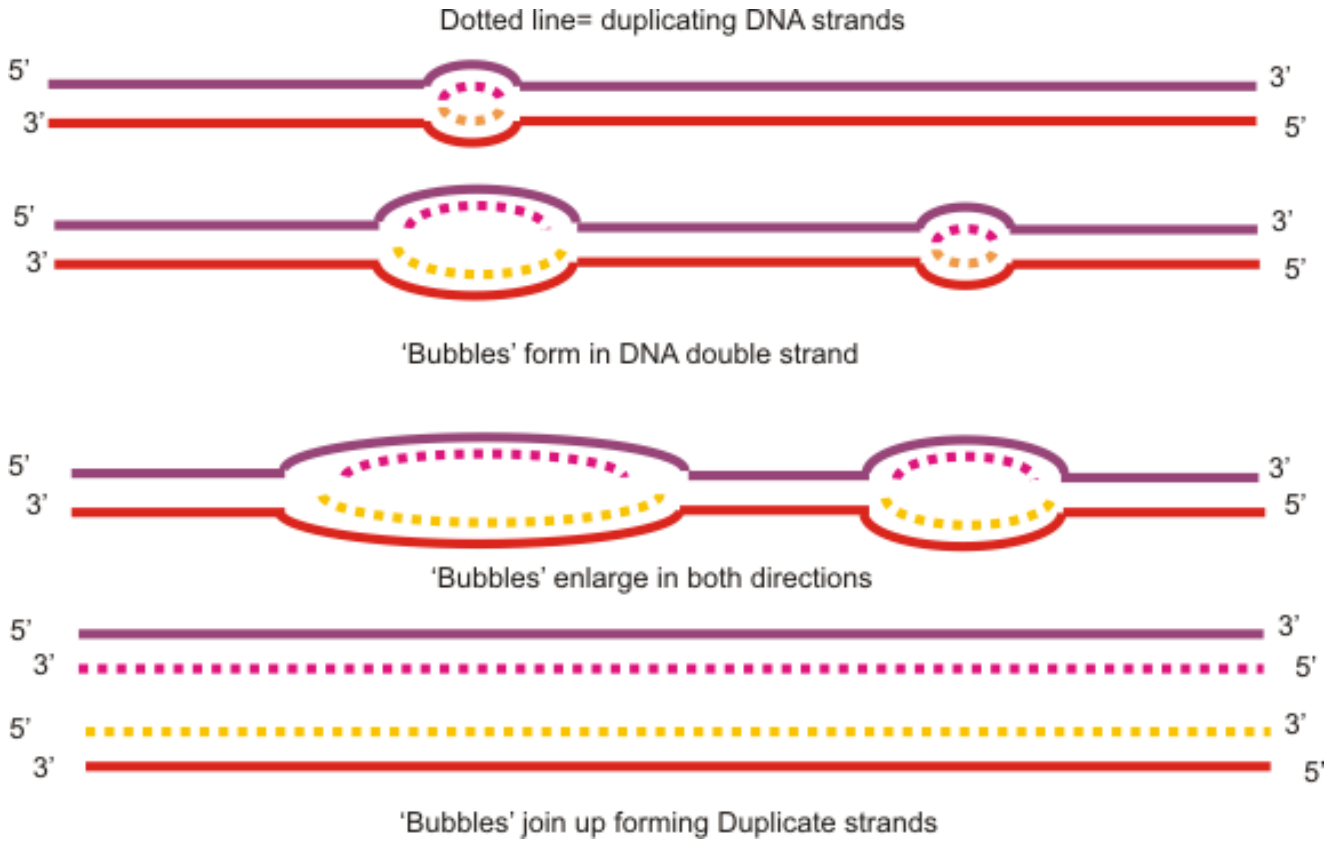


Figure 3: Multiple Replication Bubbles

DNA Repair

- Genes encode proteins that correct mistakes in DNA caused by incorrect copying during replication and environmental factors such as by-products of metabolism, exposure to ultraviolet light or mutagens.
- The DNA repair process must operate constantly to correct any damage to the DNA as soon as it occurs.

Differences between RNA & DNA

- RNA is single stranded - DNA is double stranded
- RNA has Ribose – DNA has Deoxyribose
- RNA has Uracil – DNA has Thymine

GENE EXPRESSION

Transcription and Translation utilize the DNA template code to ultimately produce proteins:

- **Transcription** – DNA is template for making RNA (in nucleus) There are 3 types of RNA.
- **Translation (protein synthesis)** - in cytoplasm at the ribosome. m-RNA has blueprint, t-RNA transfers amino acids, and Ribosome (r-RNA) allows t-RNA to attach to m-RNA at appropriate site.
- many factors control gene expression including:
  - factors affecting DNA structure,
  - gene expression,
  - factors affecting assembly of proteins after
  - translation,
  - hormones,
  - environmental factors such as viruses.

Types of RNA

Kinds of RNA – three kinds of RNA are produced in the nucleus using DNA coding strands

- **Messenger RNA (m-RNA)** – carries genetic code from DNA into cytoplasm
- **Transfer RNA (t-RNA)** – brings the amino acids for building of protein to be attached according to the genetic code of the m-RNA
- **Ribosomal RNA (r-RNA)** – make up the ribosome with a protein unit and reads the code of m-RNA and allow t-RNA to attach and connect amino acids

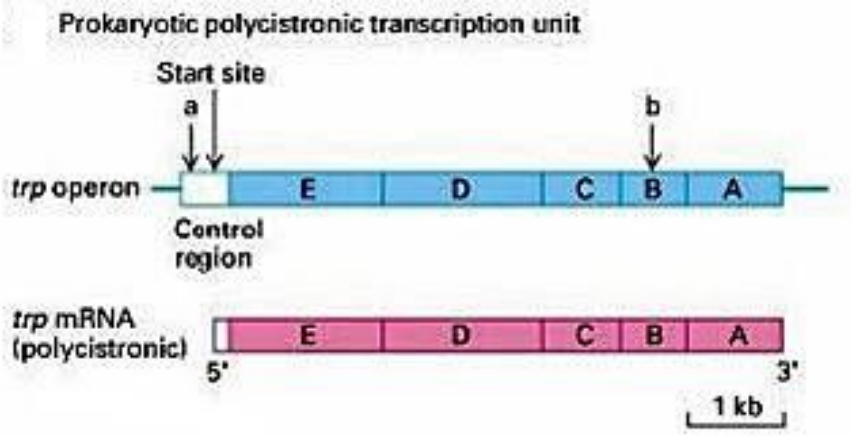
MicroRNAs*(miRNAs)*

- miRNAs are RNA genes (20-25 nucleotides long) which are transcribed from DNA, but are **not translated** into protein (non-coding RNA)
- Small non-coding **RNA** molecule which functions in transcriptional and post-transcriptional regulation of gene expression
- MicroRNAs are a class of post-transcriptional regulators
- They have the ability to regulate gene expression.
- ***MicroRNAs*** are a type of regulatory RNA that can inhibit gene expression by halting translation.
- They do so by binding to a specific location on mRNA, preventing the molecule from being translated.
- MicroRNAs have also been linked to the development of some types of cancers and a particular chromosome mutation called a translocation.

Polycistronic mRNA vs. Monocistronic mRNA

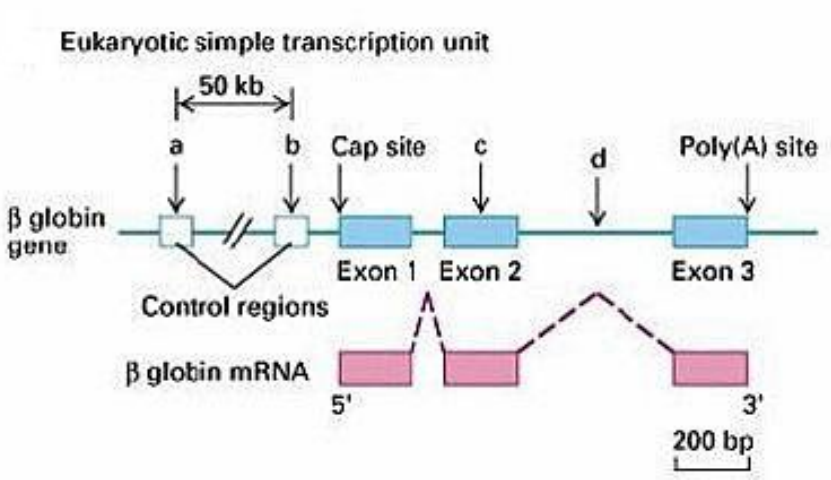
Polycistronic mRNA

- contains for more than one cistron
- codes for fore than one protein
- is transcribed from more than one gene (cistron) and has as many initiation and termination codes
- is present in **prokaryotes**



Monocistronic mRNA

- contains codons of a single cistron
- codes for a single protein
- is transcribed from a single gene (cistron) and has one initiation and termination codon
- is present in **eukaryotes**



Control of Gene Expression in Prokaryotes

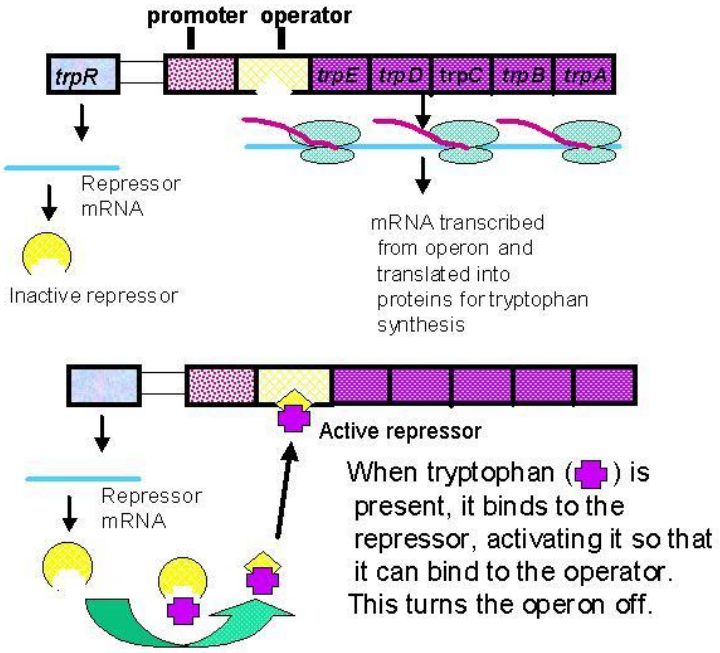
Gene expressions levels are strictly controlled at many levels to ensure the organism having the appropriate response to its environment or internal changes. This is important for prokaryotes because they are usually single-cell organisms, and they largely depend on their environment for all of their activities

In bacteria transcription often occur as polycistrons, i.e., many functional-related genes are clustered and transcribed under the same types of regulation. These are called **operons.** An operon usually contains regulatory genes and structure genes. The gene expression can be induced under certain circumstances or be constitutive.

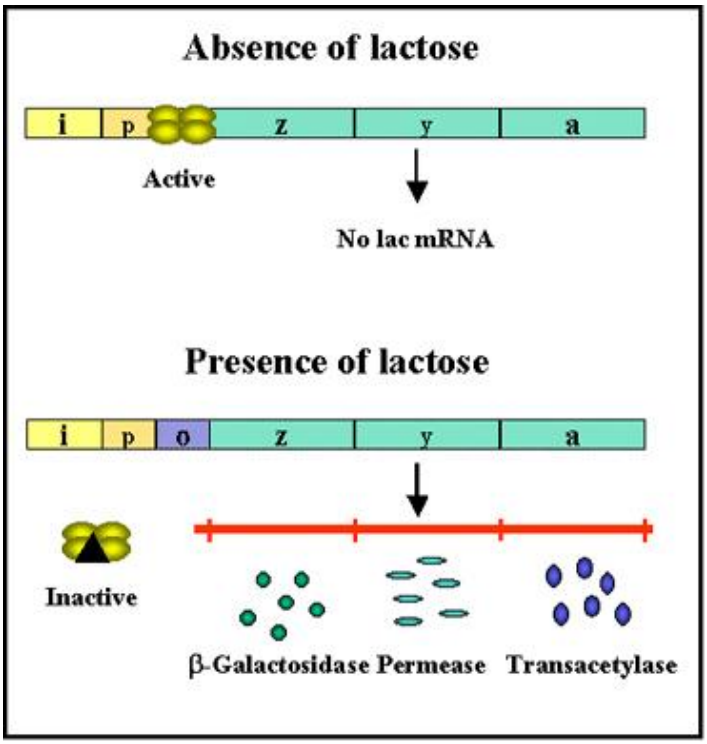
Lac & Trp Operons - *examples of prokaryotic gene regulation*

- Many of the prokaryotic genes as in *E.coli* are expressed or are always turned "on".
- Others are active only when their products are needed by the cell, so their expression must be regulated.
- Examples of Operons in *E. coli*

  - The genes for the five enzymes in the Trp synthesis pathway are clustered on the same chromosome in what is called the **Trp Operon** –
  - If the amino acid tryptophan (**Trp**) is added to a culture of *E coli* , the bacteria soon stop producing the five enzymes needed to synthesize Trp from intermediates produced during the respiration of glucos so the presence of the products of enzyme action **represses** enzyme synthesis
  - This is a **repressable operon** where genes are expressed in the absense of a substance and the presense of the substance shuts off the gene



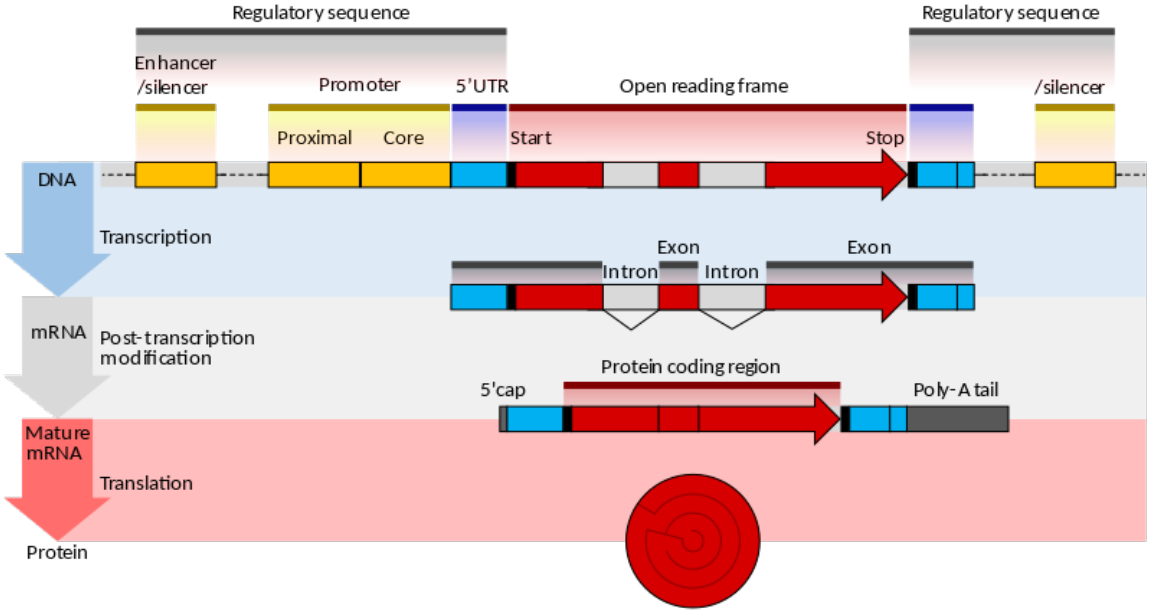
  - The genes that code for the enzymes needed for lactose catabolism are clustered on the same chromosome in what is called **the Lac Operon**
  - prokaryotes, such as **E. coli** have a mechanism for metabolizing lactose – a sugar used for energy.
  - Three proteins or enzymes are needed in lactose metabolism and they are encoded in a single expressible unit of DNA called the ***lac* operon.**
  - ***E. coli*** only express the genes and make these enzymes when lactose is available to be metabolized.
  - This is an **inducible operon** where genes are expressed in the presence of a substance



Prokaryotic gene organization

Promoter typically within 500bp of translational start site (ATG start codon of ORF)

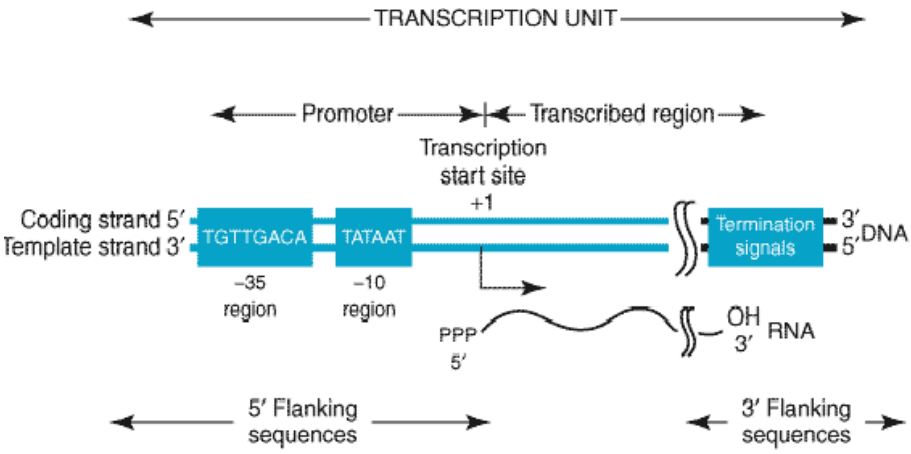
- Often (not always) polycistronic, containing more than one ORF per promoter. Co-transcribed genes are said to be in an operon.
- No introns are present
- Multiple ribosome binding site (RBS) present in polycistronic mRNA



- Extended promoter with enhancers/silencers often more than 1kb away
- RNA is processed (post-transcription modifications)
  - Introns are excised and exons spliced together
  - 5' cap and poly-A tail added to transcripts
- Only one open reading frame (no operons/polycistronic mRNAs)
- No ribosome binding site on mRNA
  - ribosome binds 5' end of mRNA and scans for translational start codon

Prokaryotic promoter - region of DNA that initiates transcription of a particular gene

- The -10 and -35 regions of a canonical prokaryotic promoter are recognized by the bacterial RNA polymerase sigma subunit.
- The site of transcription initiation occurs 10bp downstream of the -10 region's center, hence its name. The consensus sequences for the -35 and -10 sequences are shown.
- These sequences lead to strong transcription of a promoter, but variations of the sequences are common and can still interact strongly with the RNA polymerase sigma subunit.



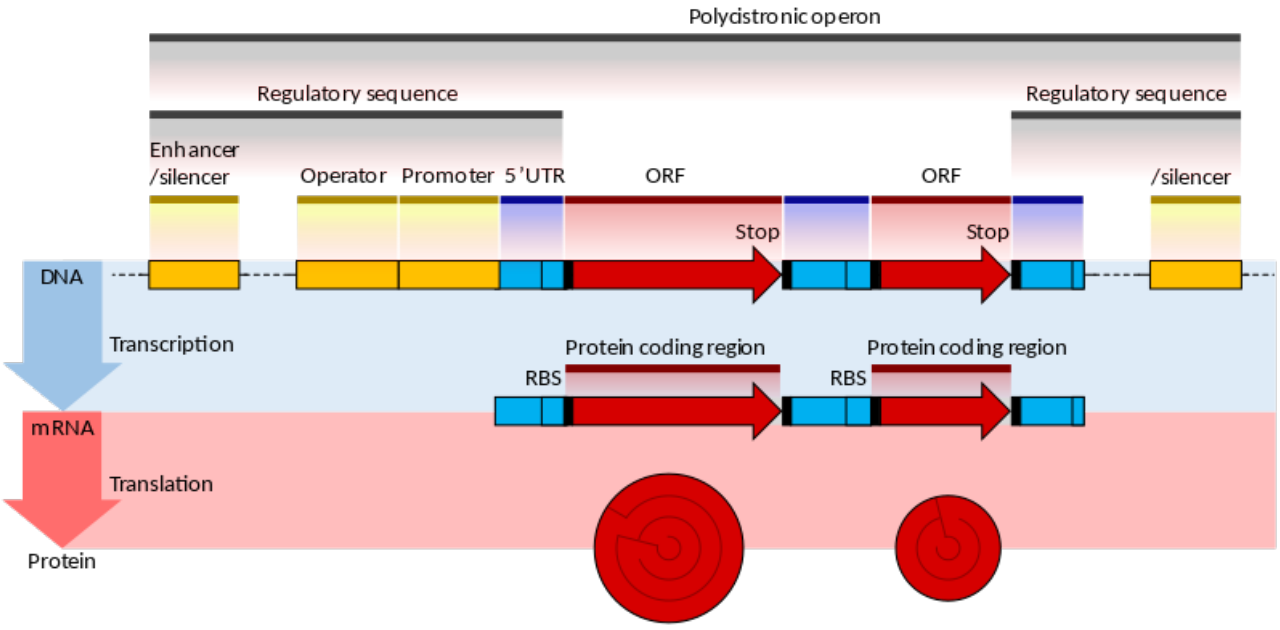
Control of Gene Expression in Eukaryotes

Eukaryotic genes usually contain three basic regulatory components:

- **Enhancers** - short regions of DNA that can be bound with proteins to *promote expression* of a distal or a proximal gene.
- **Promoters** - proximal DNA sequences that binds to RNA polymerase *for regulating gene expression*.
  - o region of DNA that initiates transcription of a particular gene
  - o located near the genes they transcribe, on the same strand and upstream on the DNA (towards the 3' region of the anti-sense strand also called template strand and non-coding strand

- **TATA Box** - binds to transcription *factor for regulating gene expression*, usually within 30bp of the transcription start site.

Components of a Genes and promoters

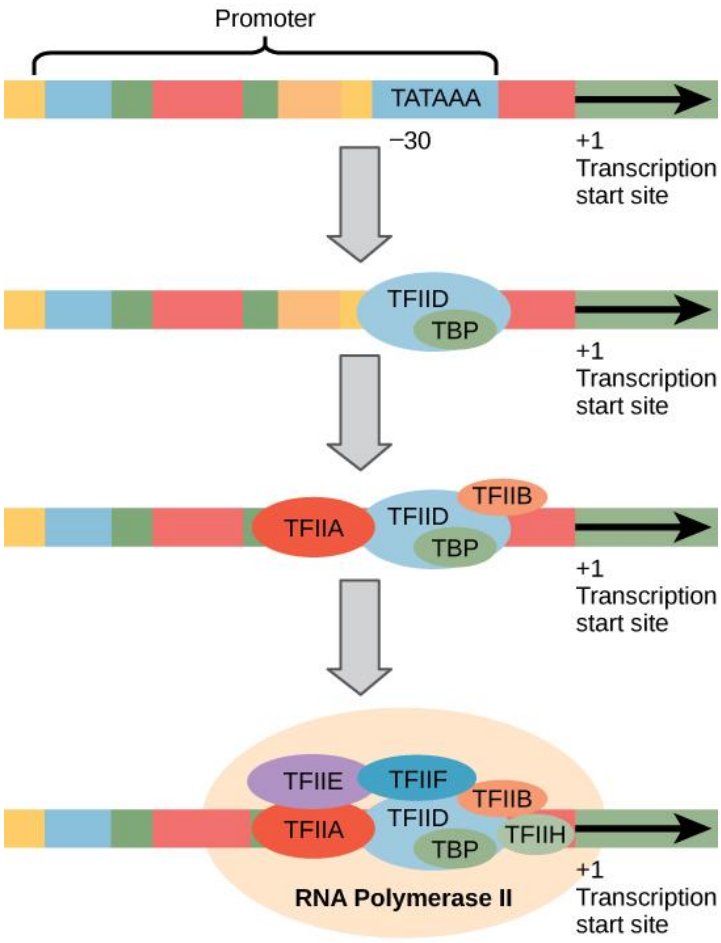


- **Promoter**: regulatory sequence upstream of gene to which RNA polymerase binds to initiate transcription
- **Transcriptional start site**: nucleotide within promoter where transcription begins. It is upstream of the ATG start codon where translation begins. Difficult to identify with genome scanning software, so often needs to be determined experimentally.
- **5' Untranslated Region (UTR)**: stretch of mRNA between the transcriptional start site and the translational start site, often controls rate of translation initiation
- **Open Reading Frame (ORF):** the part of a gene that codes for protein. Begins with an ATG start codon and ends with a stop codon that are in frame with each other. Very easy to recognize with genome scanning software.

- **3' Untranslated Region (UTR):** stretch of mRNA between the final stop codon and the end of the transcript. Like 5' UTR, 3' UTR can also be regulatory.

Eukaryotic promoter

- Eukaryotic promoters also contain a "TATA" box upstream of the transcriptional start site, and this sequence is bound by a series of transcription factors that recruit RNA polymerase to initiate basal levels of transcription.
- Like in prokaryotes, the rate of transcription is then further influenced by other transcriptional activators/repressors that bind other locations in the promoter and/or at enhancer sequences.
- Unlike prokaryotes, these enhancer sequences can often be very distant from the gene they regulate. DNA with the enhancer sequence can loop back to allow a transcriptional regulator to interact with RNA polymerase at the promoter.

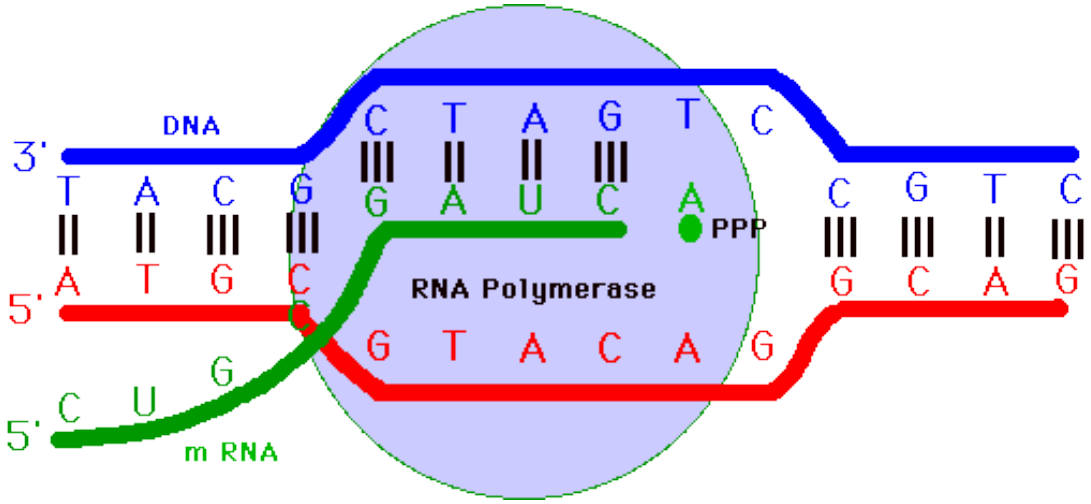


Controls include:

- Transcriptional Control
- Post transcriptional Control – assembling proteins
- Cell differentiation and specialization
- Turning genes "on" and "off"
- Chemical Signals – Hormones
- Chemical Modifications
- Relocation of DNA – transposons
- Abnormal Expression of Genes

Transcription

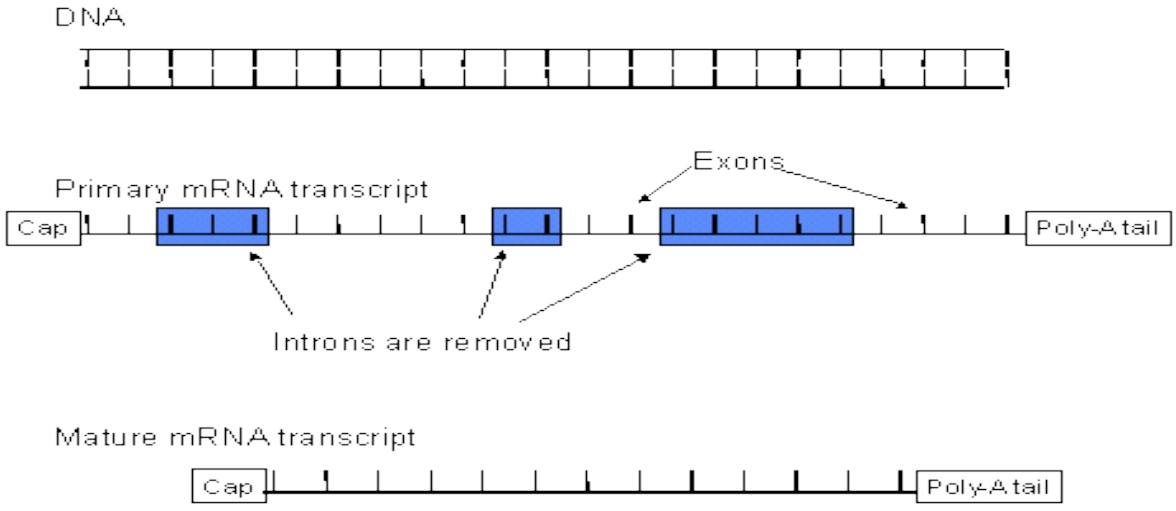
Transcription - production of RNA in the nucleus using a DNA segment as a template and **RNA polymerase** as the key enzyme.



Post-transcription Modifications in Eukaryotic Cells

RNA's are modified in eukaryotes before leaving the nucleus.

- Pre-m-RNA has **exons** (coding segments) and **introns** (noncoding segments between exons)
- **introns** (the noncoding segments) are removed
- a cap is added to the 5' end
- a poly A tail is added to the 3' end before it leaves the nucleus



Universal Code (Codon = Amino Acid)

First position (5' end)	U	C	A	G	Third position (3' end)
U	UUU Phe UUC UUA Leu UUG	UCU Ser UCC UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trp	U C A G
C	CUU Leu CUC CUA CUG	CCU Pro CCC CCA CCG	CAU His CAC CAA Gln CAG	CGU Arg CGC CGA CGG	U C A G
A	AUU Ile AUC AUA AUG	ACU Thr ACC ACA ACG	AAU Asn AAC AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
G	GUU Val GUC GUA GUG	GCU Ala GCC GCA GCG	GAU Asp GAC GAA Glu GAG	GGU Gly GGC GGA GGG	U C A G

Amino acid names:

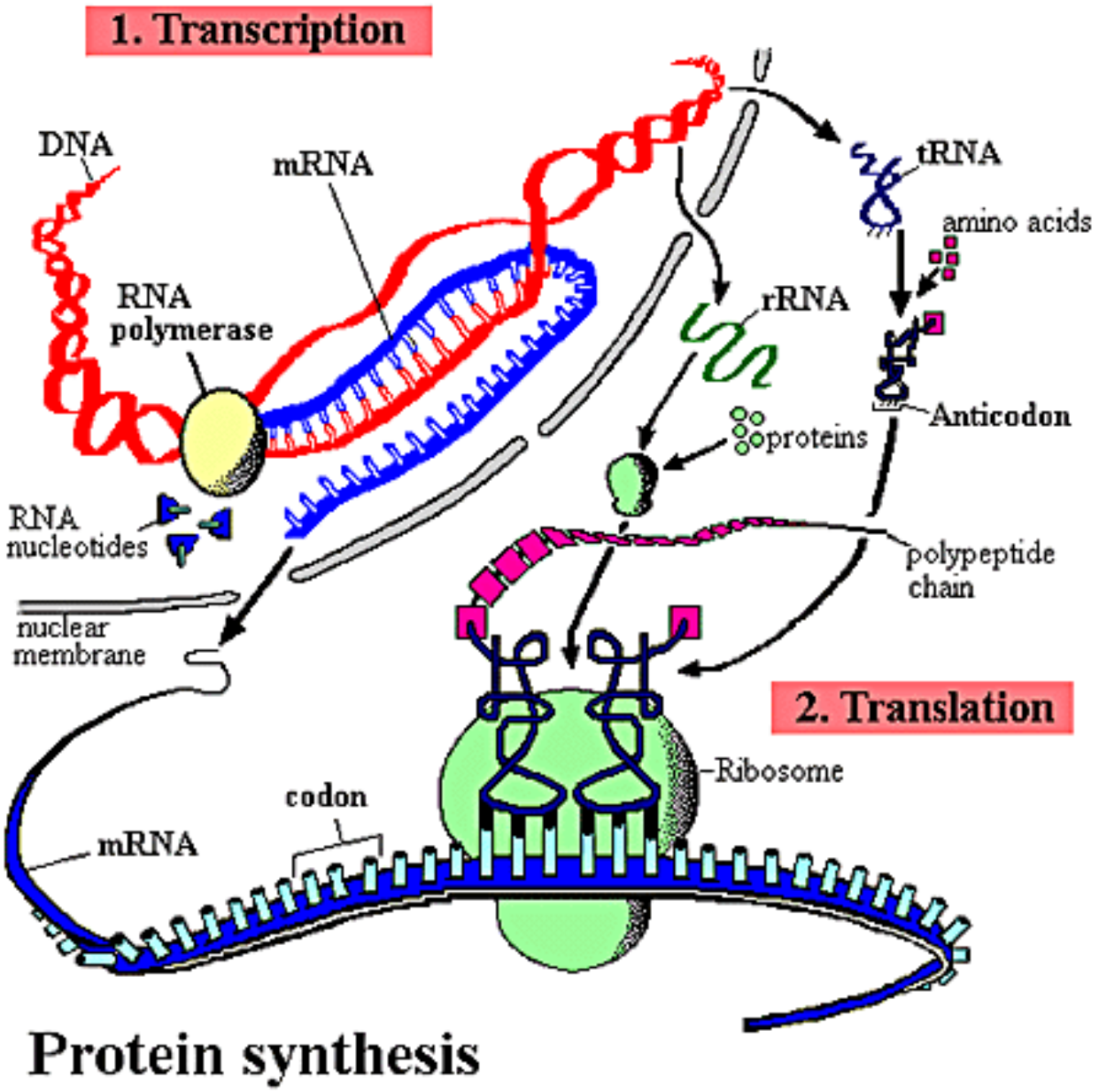
Ala = alanine	Gln = glutamine	Leu = leucine	Ser = serine
Arg = arginine	Glu = glutamate	Lys = lysine	Thr = threonine
Asn = asparagine	Gly = glycine	Met = methionine	Trp = tryptophan
Asp = aspartate	His = histidine	Phe = phenylalanine	Tyr = Tyrosine
Cys = cysteine	Ile = Isolevcine	Pro = proline	Val = valine

- Each three-base codon on the messenger RNA (m-RNA) is a code for an amino acid
- There are 64 possible three base codons – 61 are codes for one of the 20 amino acids
- The three remaining codons are termed **stop codons** because the signal the end of a peptide segment
- Notice that many of the amino acids have more than one codon
- A three-base **code** on the **DNA** produces the **mRNA codon**
- The three-base code on the **t RNA** is termed an **anticodon** because it will bond to a m-RNA codon during translation or protein synthesis

Translation (Protein Synthesis)

- Translation is the universal process of synthesizing proteins as the second step in gene expression.
- Both prokaryotic and eukaryotic ribosomes decode mRNAs in fundamentally similar methods
- The main difference between prokaryotic and eukaryotic translation is that prokaryotic translation is a simultaneous process with transcription whereas eukaryotic translation is a separate process from its transcription.
- **Translation** – m-RNA template containing genetic code is used to form amino acid sequence using m-RNA, t-RNA, and r-RNA (ribosomes) occurs in the cytoplasm at the ribosome. Many key enzymes (proteins) are involved.

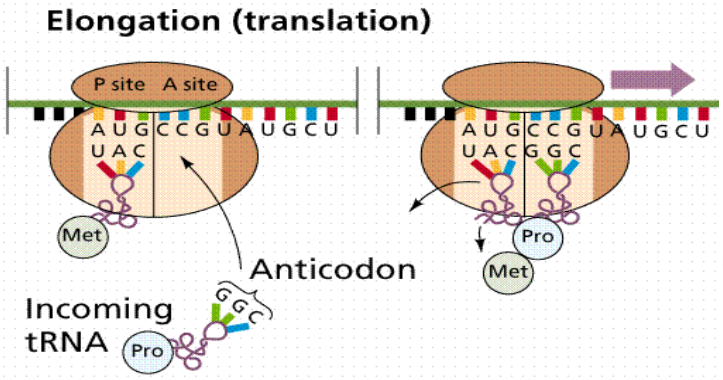
.

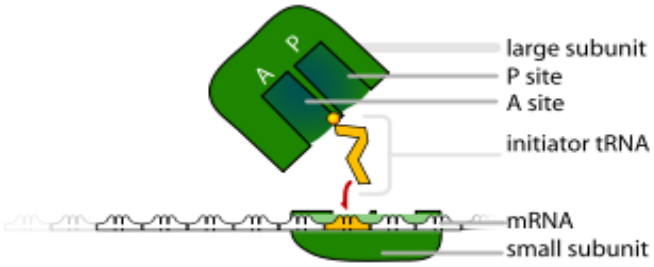


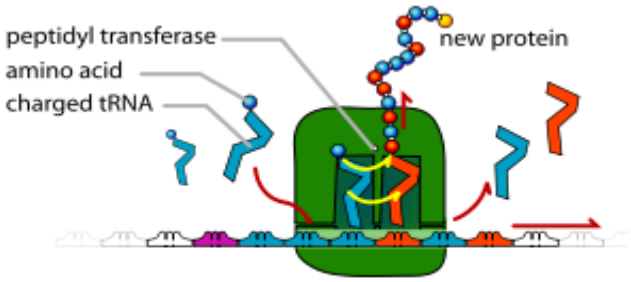
Protein synthesis

The steps of translation:

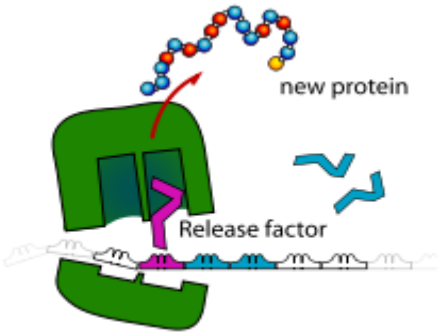
- **Initiation**: **mRNA** enters the cytoplasm and becomes associated with **ribosomes** (**rRNA** + **proteins**) and **tRNAs**, each carrying a specific amino acid, pair up with the mRNA codons inside the ribosomes. The base pairing (A-U, G-C) between **mRNA codons** and **tRNA anticodons** determines the order of amino acids in a protein.
- **Elongation:** involves the **addition of amino acids one-by-one**: As the **ribosome** moves along the **mRNA**, each **tRNA** transfers its amino acid to the growing protein chain, producing the protein







A close up showing the M-RNA (with its codon) and T-RNA (with it anticodon as well as the Amino Acid) attaching at the P and A sites on the Ribosome.



- **Termination:** when the **ribosomes** hits a stop codon - UAA, UGA, or UAG – no tRNA with its amino acid can be added so the ribosome falls apart and the process ends. The same **mRNA** may be used hundreds of times during translation by many **ribosomes** before it is degraded (broken down) by the cell.

Nuclear vs Cytoplasmic DNA in Eukaryotic Cells

- **Nuclear DNA** – in chromosomes within the nucleus of the cell
- **Cytoplasmic (or Organelle DNA)** – in chloroplasts and mitochondria
    - Mitochondria and Chloroplasts have DNA similar to Prokaryotic cells
    - It is believed that these organelles were once independent prokaryotes who took up residence and formed a mutualistic relationship
    - They are involved in energy transfer- photosynthesis & respiration

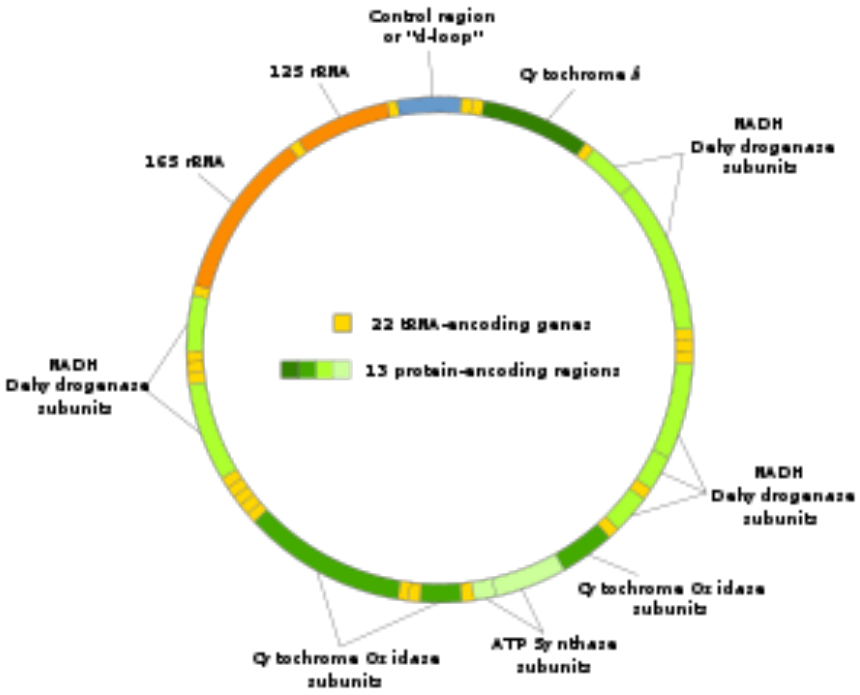
    - **Chloroplast DNA (cpDNA)**
    - **Mitochondrial DNA (mtDNA)**

      Features:

        - Maternal inheritance
        - Resemble prokaryotic DNA
        - Slow accumulation of mutations at the population level

Mitochondrial Inheritance –

- The **inheritance** of a trait encoded in the **mitochondrial** genome
- Mitochondrial DNA or mtDNA -genetic make-up of mitochondria, genetic code and patterns transmitted through mother.
- The mtDNA is circular and resembles prokaryotic DNA
- The mitochondria are responsible for energy production
- Mitochondria can reproduce independent of the rest of the cell – an advantage in energy production
- Persons with a mitochondrial disease may be male or female but they are always related in the maternal line and no male with the disease can transmit it to his children
- Mitochondrial myopathies are a group of neuromuscular diseases caused by damage to the mitochondria-small, energy-producing structures that serve as the cells' "power plants."



Mutations

- Gene – section of DNA which carries the blueprint for making a peptide strand or RNA.
- DNA in the living cell is subject to **many chemical alterations** - If the genetic information encoded in the DNA is to remain uncorrupted, any chemical changes must be corrected.
- A failure to repair DNA produces a **mutation**
- Mutation – changes in genetic code (DNA blueprint) of genes or chromosomes and causes changes in expression or in the directions for making protein or RNA
- Gene mutation
- Chromosomal mutation
- Agents causing mutations – radiation, chemicals, excess heat , viruses

Mutagenesis is the process of making mutations. These mutations can be random or can target specific genes.

Forward genetics is the classical approach to genetics where mutations are randomly made in the genome and the experimenter screens for a desired phenotype. Mutations that lead to the desired phenotype are then mapped to determine which gene(s) were mutated.

Common examples of random mutagenesis: radiation, chemicals, transposon insertions

Reverse genetics is a more recent approach to genetics in which a specific gene of interest is mutated in order to determine the phenotype(s) that arise as a result of the mutation.

Common examples of targeted mutagenesis: targeted transposon insertion, deletion, allelic exchange, CRISPR-Cas modifications

Genetic Disorders

- **Causes of mutations –** chemicals, radiation, temperature, viruses
- **Nondisjunction** – chromatids do not separate properly during meiosis. Individual formed from such gametes have extra or missing chromosomes. as Down's Syndrome
- **Trinucleotide repeats –** sequences of 3 nucleotides is repeated, often several times in a gene. When too many repeats are formed it cause genetic disorders from the triplet nucleotides being repeated too often, such as Huntington's
- **Defective genes** – does not produce correct protein, such as sickle cell anemia (A & T traded places)
- **Genetic disorders and their causes** as nondisjunction (Down's syndrome), trinucleotide repeats (fragile X and Huntington's), defective genes (sickle cell anemia, hemophilia)
- **Human genetic disorders –** can be dominant, recessive, sex-linked, epistatic, variable expressed
- **Crossover frequency –** during meiosis, pieces trade places – determining frequency

Hardy-Weinberg Equilibrium

Hardy-Weinberg Equilibrium is a theorem that describes allele and phenotype frequencies in a large population without gene flow/introduction. A population in Hardy-Weinberg equilibrium will have the following primary assumptions:

- **Random mating-** There is no sexual selection
- **Infinite population size-** Eliminates the effect of genetic drift
- **No migration-** Additional alleles are not being introduced from the outside
- **No mutation-** Additional alleles are not popping up spontaneously
- **No selection-** One allele is not selected for over another, keeping gene frequencies constant

These conditions will never happen in nature, so actual populations are never in Hardy-Weinberg equilibrium. However, it does provide a baseline to compare actual populations to and see how they change over time.

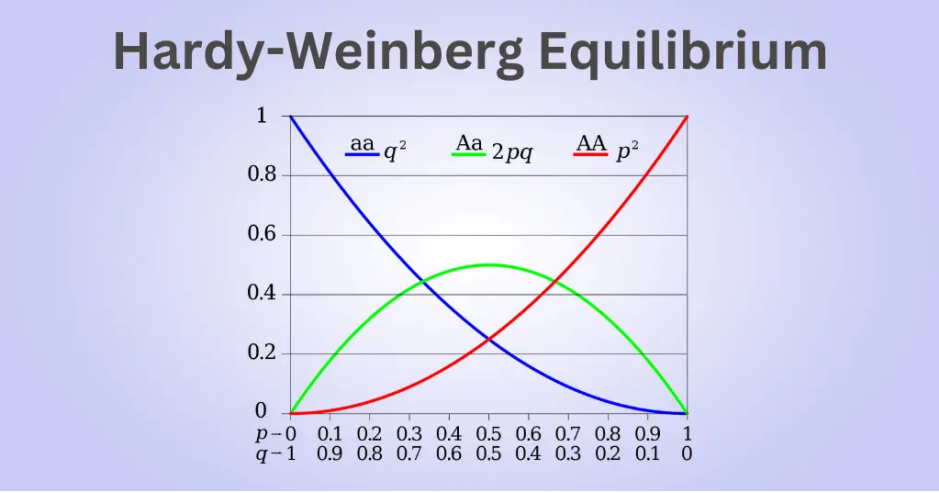
The equations describing a population in Hardy-Weinberg Equilibrium are:

$$p + q = 1$$

This equation states that the frequencies of allele p and q must equal 1.

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

This equation states that the frequencies of individuals homozygous for the dominant allele (shown by p^2), heterozygotes (2pq), and homozygous for the recessive allele (shown by q^2) also must equal 1. If you know the frequency of individuals with a certain genotype, you can calculate the frequencies in general or refer to the chart below.



For a more fun explanation, here's a video: https://www.youtube.com/watch?v=7S4WMwesMts