Fundamentals of Human Genetics

JONATHAN POHL, MLS (ASCP)^{CM} MB^{CM}

CLEVELAND CLINIC CENTER FOR PATHOLOGY EDUCATION

Disclaimer

This presentation was developed by Cleveland Clinic's Center for Pathology Education (CPE) for use in Cleveland Clinic's Molecular Biology Tech Categorical Exam Preparation Program (TCEPP) and Cleveland Clinic's School of Medical Laboratory Science.

This presentation is intended for educational purposes only.

Information in this presentation is not meant to substitute or replace the policies and standard operating procedures of Cleveland Clinic or RT-PLMI.

Images of and references to specific brands of instrumentation, consumables, reagents, or other purchasable items are used as examples and are not endorsements of those specific brands or items.

Objectives

- Define and differentiate fundamental concepts in human genetics:
  - Gene/Locus/Allele
  - Genotype/Phenotype
  - Genome/Exome/Transcriptome/Proteome
  - Types of polymorphisms and mutations
  - Types of genetic disorders
  - Patterns of inheritance
  - Epigenetic factors

References and Additional Resources

- Buckingham, Lela. *Molecular Diagnostics: Fundamentals, Methods, and Clinical Applications.* 3rd ed., F.A. Davis, 2019.
  - Chapter 1: Nucleic Acids and Proteins
  - Chapter 2: Gene Expression and Epigenetics
  - Chapter 12: Molecular Detection of Inherited Diseases (pages 344-349)

- Nussbaum, Robert, et al. *Thompson & Thompson Genetics in Medicine*. 8th ed., Elsevier, 2016
  - Chapter 7: Patterns of Single-Gene Inheritance
  - Accessible free online at https://www.clinicalkey.com/#!/browse/book/3-s2.0-C2009059798X

What is a gene?

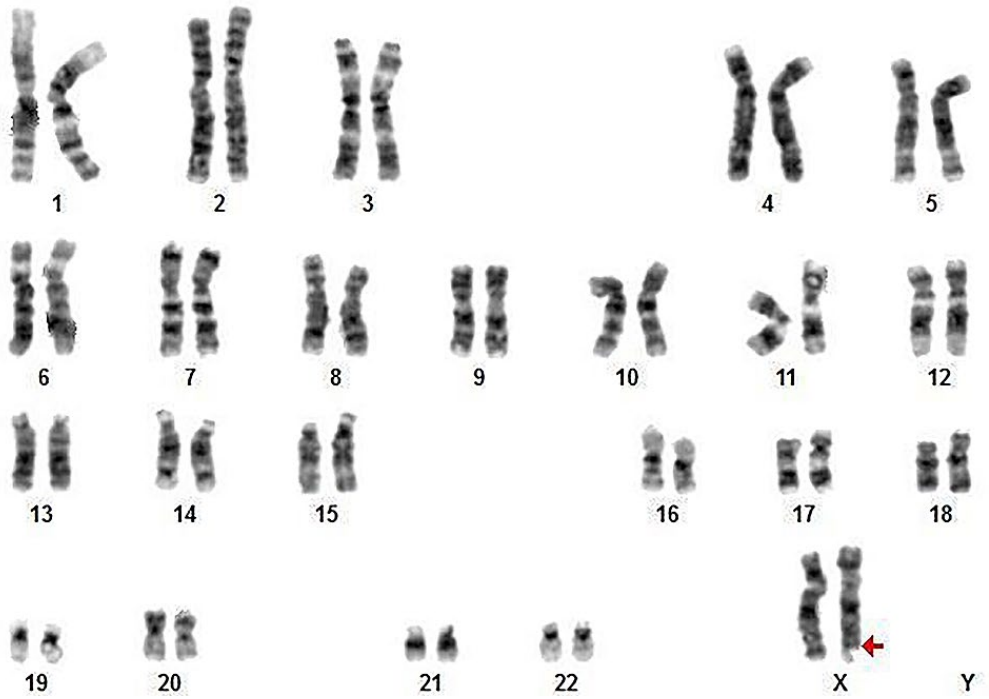
A **gene** is a sequence of nucleotides that contains all the genetic information to make a functional protein or RNA product.

A **locus** (plural, **loci**) is the specific location of a gene within an organism's genome.

An **allele** is a version of a gene.

For example…

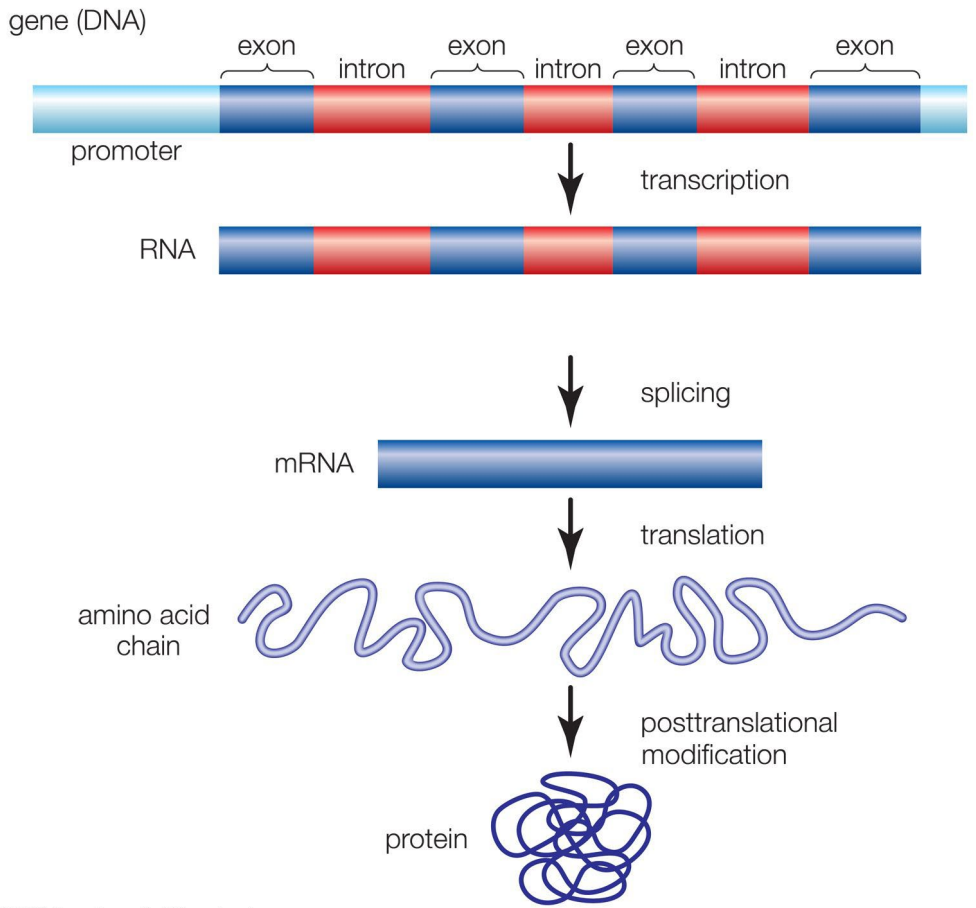
- FMR1 is a gene that codes for the FMRP protein.
- The FMR1 gene locus exists on the q-arm of the X chromosome.
- This means female (46,XX) patients will have two alleles capable of producing functional FMRP, while male (46,XY) patients will have one allele.



Gene Structure

Gene sequences are composed of three major sections:

- **Promoter** = DNA sequence that binds RNA polymerase for mRNA transcription
  - Situated “upstream”, or 5’, of the coding region of the gene
- **Exons** = DNA sequences that code for protein
- **Introns** = noncoding sequences that interrupt exonic sequences, which are removed post-transcription in a process called **splicing.**
  - **Alternative splicing** of introns allows a single gene to generate multiple products.



Genome / Exome / Transcriptome / Proteome

Genome: the complete set of genes/genetic material of an organism

- Includes both exonic and intronic sequences

Exome: the complete set of exonic sequences of an organism

- Does not include introns

Transcriptome: the complete set of mRNA expressed by an organism

Proteome: the complete set of proteins produced of an organism

Variation in the Human Genome

When comparing any two random human genomes, approximately 99.6% of the sequences will be identical.

The 0.4% of variable genomic material is what accounts for the wide range of phenotypic variation between individuals.

- **Genotype**:
  - In the book: the genetic DNA composition of an organism
  - In the lab: the specific allelic arrangement of a particular gene (*e.g.*, HET, HOM, WT)

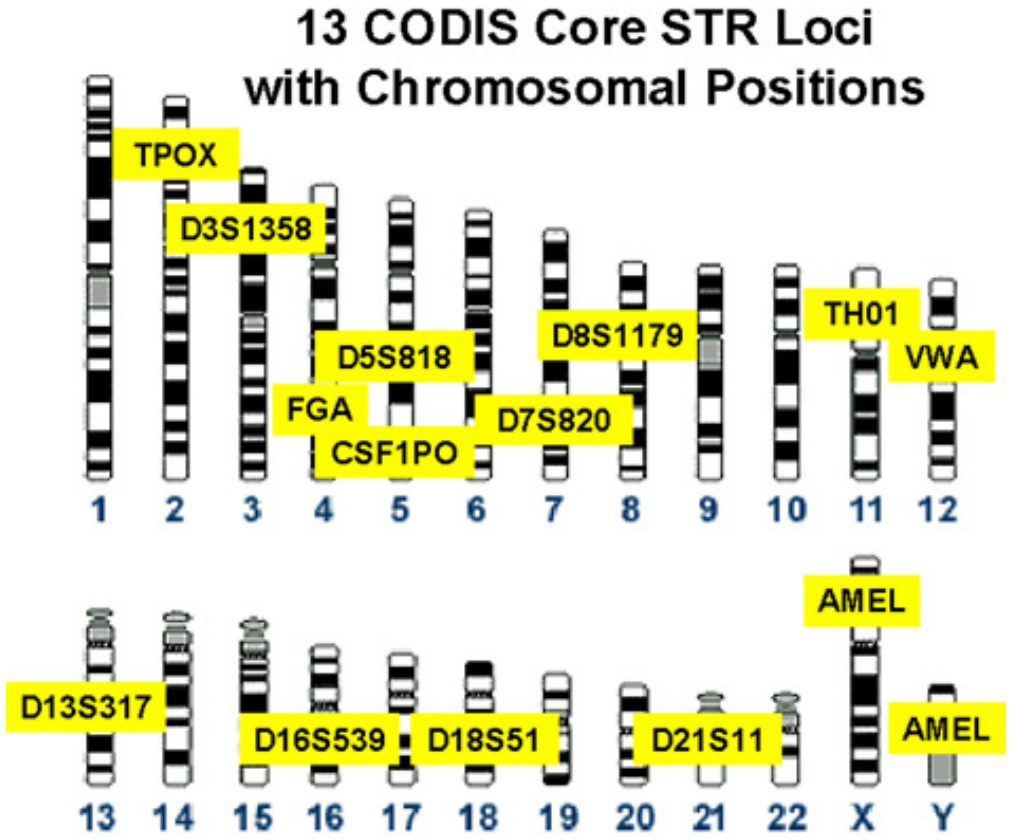
- **Phenotype**:
  - In the book: the biological properties of an organism
  - In the lab: the biological properties produced by a specific genotype (*e.g.*, rapid, intermediate, or low metabolizer of a drug)

Polymorphism

Polymorphism represents a genetic variation within a population at levels greater than 1%.

This distinguishes it from rare variants or de novo mutations.

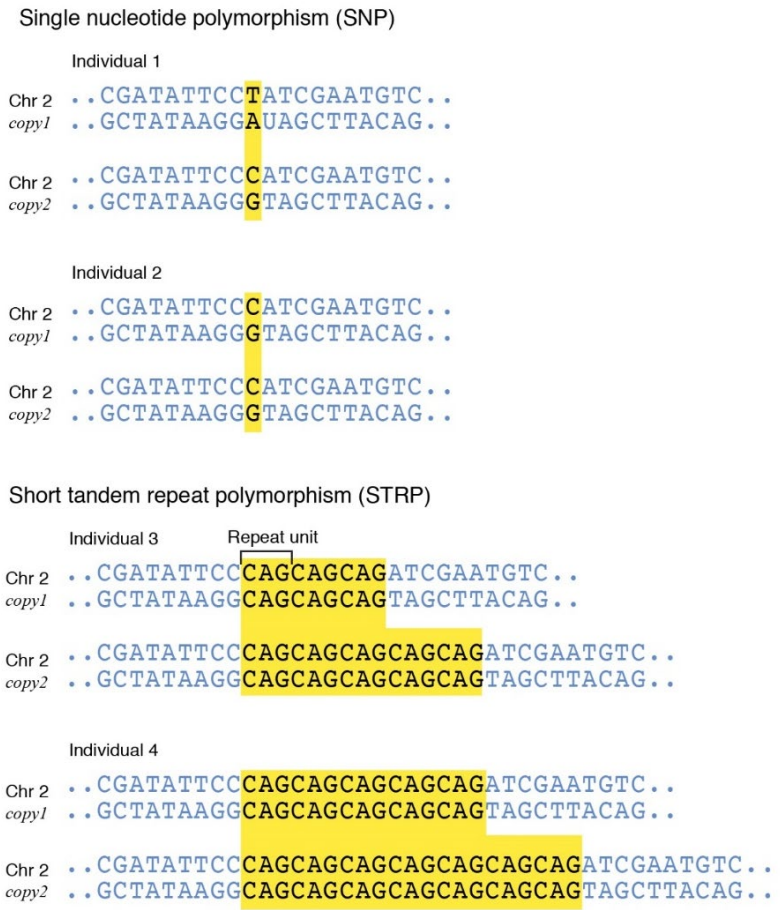
Polymorphisms may or may not cause disease/correspond with any specific phenotypic presentation.



Polymorphism

Common types of genetic polymorphism:

- **Single nucleotide polymorphism** (**SNP**) = single base pair variance, also called single nucleotide variant (SNV)
- **Copy number variant** (**CNV**) = whole gene(s) or chromosomal regions that are duplicated/triplicated/etc.
- **Short tandem repeat** (**STR**) = >1 but <10 base pair sequence that repeats itself
- **Variable number tandem repeat** (**VNTR**) = >10 but <50 base pair sequence that repeats itself



Mutations

The words **variant** and **mutation** are used mostly interchangeably to refer to an alteration in the sequence of a specific gene.

- Variant = usually a **heritable** change in the gene, may or may not cause disease (*e.g.*, most people have a variant allele in their MTHFR gene)
- Mutation = may be inherited or de novo, generally always causes disease (*e.g.*, SEA mutation for alpha-thalassemia)

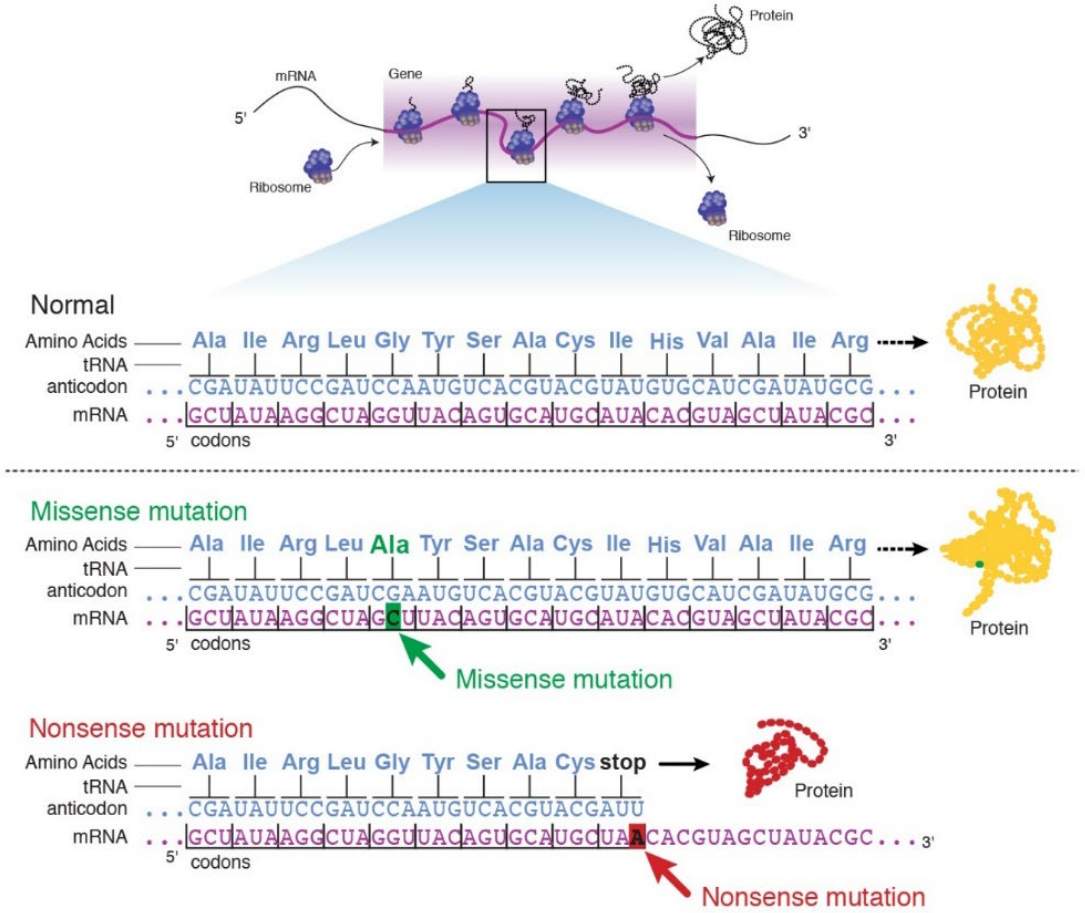
There are several kinds of mutations:

- Point Mutations
- Insertions
- Deletions
- Fusions

Point Mutations

Point mutation = a change of one base pair

- **Substitution** = one base pair is replaced by another
  - **Missense** = point mutation results in translation of a different amino acid
  - **Nonsense** = point mutation changes a sense codon into a stop codon
  - **Silent** = point mutation does not have an effect on protein synthesis
- **Transition** = purine is replaced with purine (A to G), or a pyrimidine with pyrimidine (C to T)
- **Transversion** = purine is replaced with pyrimidine (e.g. A to T) or pyrimidine is replaced with purine (C to G)



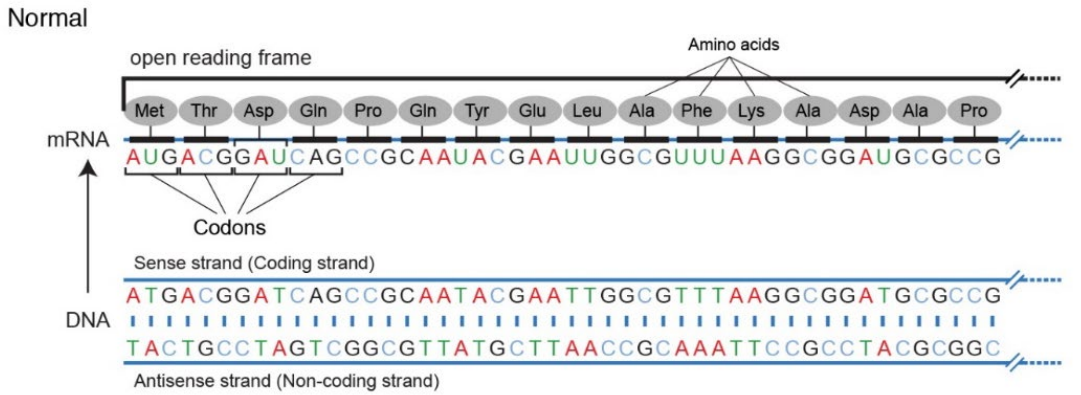
Indels

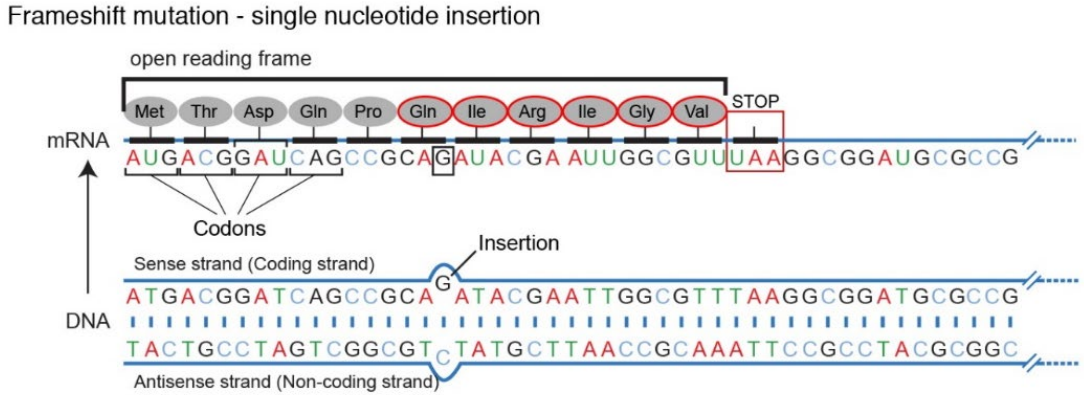
Insertions and deletions are usually collectively referred to by their portmanteau, **indels**.

- **Insertion** = base pair(s) are added to the sequence
- **Deletion** = base pair(s) are removed from the sequence

An indel that is not divisible by three is referred to as a **frameshift mutation**.

- Recall that codons are made up of 3 base pairs



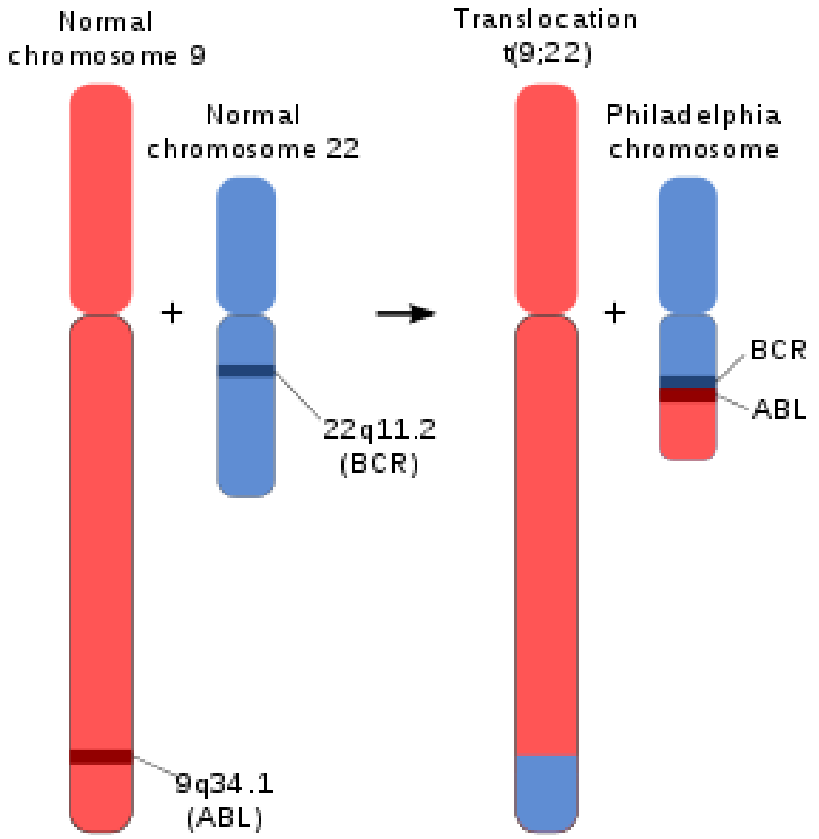


Fusions

Gene **fusions** occur when deletions, duplications, inversions, and translocations place two normally separate genes next to one another.

- **Inversion** – portion of a chromosome has broken off, turned upside down, and reattached
- **Translocation** – genetic material is transferred from one chromosome to another

The gene fusion, or **chimeric** gene, has different and often malignant activity than its original parts.



Nondisjunction

Nondisjunction is the abnormal separation of chromosomes during cell division which results in both of a chromosome pair in one daughter cell.

- Meiotic nondisjunction is the most common mutational mechanism, responsible for chromosomal disorders and gene rearrangements.
- Mitotic nondisjunction in somatic cells can cause chromosomally abnormal tumors.

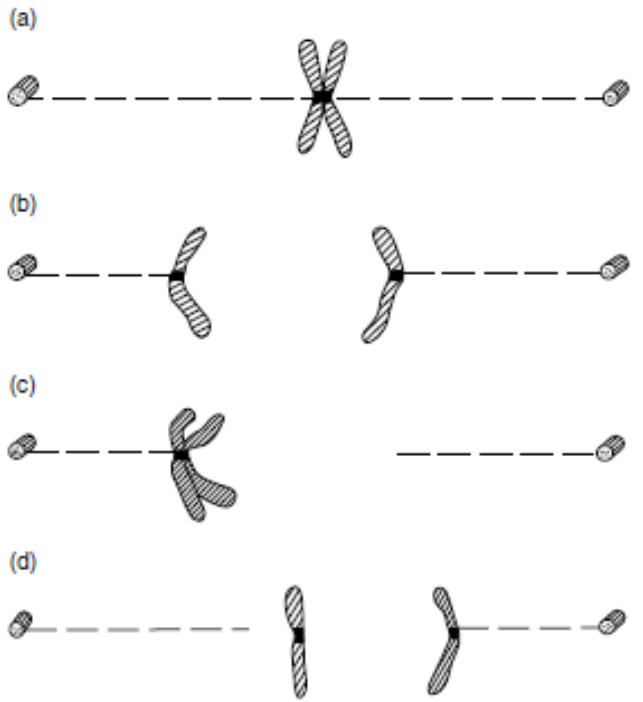


Figure 2.19 Normal and abnormal chromatid separation at mitosis. (a) Chromosomes at metaphase plate between centrioles. (b) Normal disjunction at anaphase – one chromosome at each centriole. (c) Nondisjunction at anaphase – two chromatids travel to one centriole, resulting in a cell with an extra chromosome and a cell with a missing chromosome. (d) Anaphase lag – one chromatid fails to attach to the spindle and is usually excluded from the nuclei and lost from both daughter cells.

Describing Genetic Disorders

Hereditary, or **inherited**, disorders are those passed from parent to offspring.

- *e.g.,* cystic fibrosis

Constitutional, or **germline**, disorders are those that develop prior to birth.

- Can be inherited (see above), or can result from changes to the genetic code during gamete production or embryological development (*e.g.*, trisomies)

Congenital disorders are those *present* at birth

- Can be used to describe the above scenarios but does not necessarily have to be genetic (*e.g.,* fetal alcohol syndrome).

Acquired, or **somatic**, disorders are those that develop over the course of a person's life, often as a result of damage to the genetic code.

- *e.g.,* cancer

Describing Genetic Disorders

Single-gene disorders = disorders resulting from a single affected gene

- Homozygous inheritance of the deltaF508 variant of the CFTR1 gene results in symptomatic cystic fibrosis.
- Analyzing only the CFTR gene is sufficient to predict patient outcomes, carrier status.

Polygenic disorders = disorders that result from the combined effects of multiple genes

- Another term used to describe these disorders is “multifactorial”
- Factor V and Factor II of the coagulation cascade are produced by different genes.
- Both genes are analyzed in a hypercoagulation panel to assess a patient’s overall risk for thrombosis.

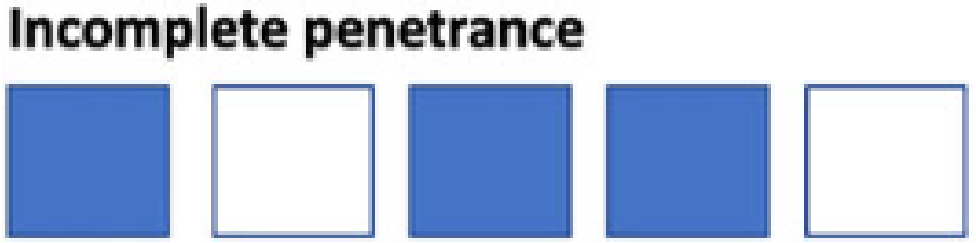
Chromosomal disorders = disorders resulting from numerical or structural changes to patient chromosomes

- Trisomy 21 (Down Syndrome) results from the numerical addition of an extra chromosome 21.

Describing Genetic Disorders

Penetrance refers to the proportion of patients with a disease-causing gene variant that present with disease.

- If a gene variant always causes disease when inherited, it is said to have **complete** (100%) penetrance.
- If a gene variant does not always cause disease when inherited, it is said to have **incomplete** (<100%) penetrance.

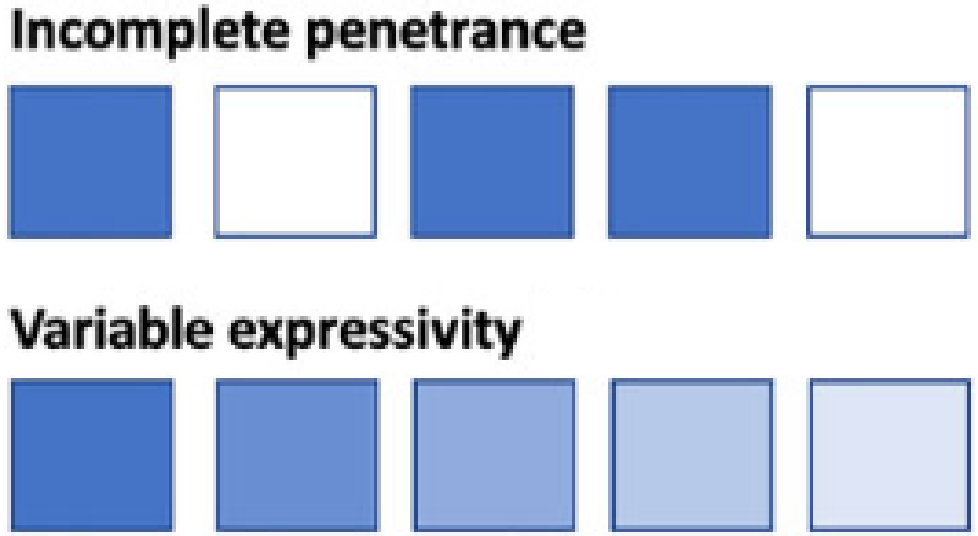


Example: Huntington's Disease

Describing Genetic Disorders

Expressivity refers to the severity of disease and range of clinical manifestations.

- Interaction with other genes, epigenetic and environmental factors can alter expressivity.

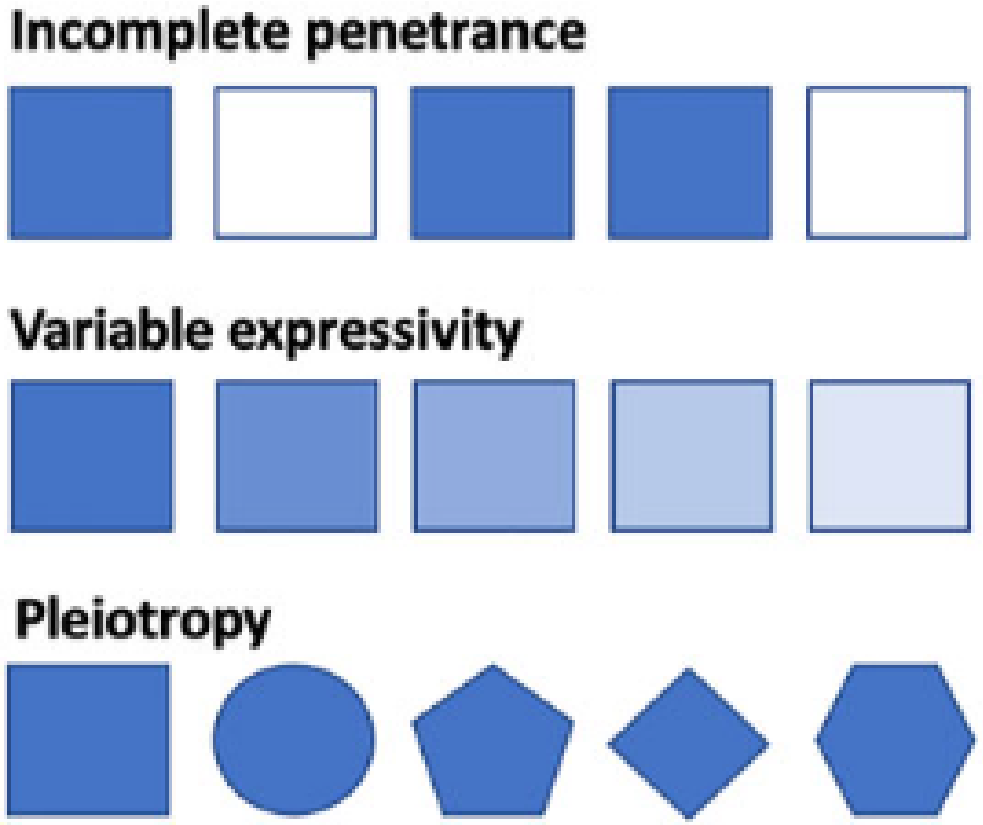


Example: Spinal Muscular Atrophy

Describing Genetic Disorders

Pleiotropy is when a single gene affects two or more phenotypic traits.

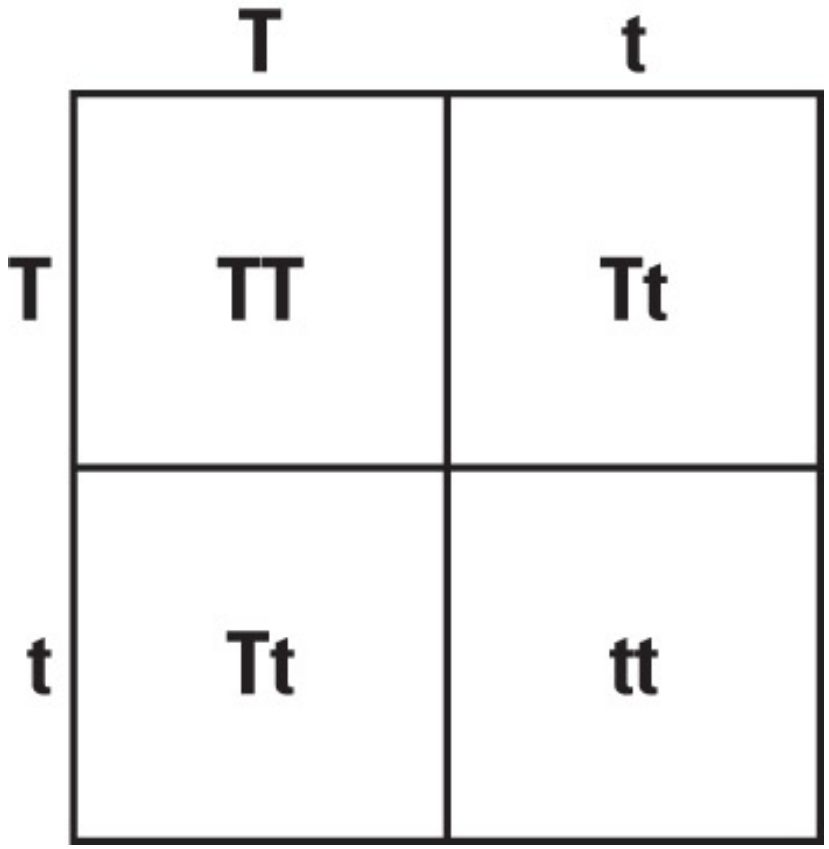
Example: Cystic Fibrosis



Describing Genotype

Describing genotype for single-gene disorders:

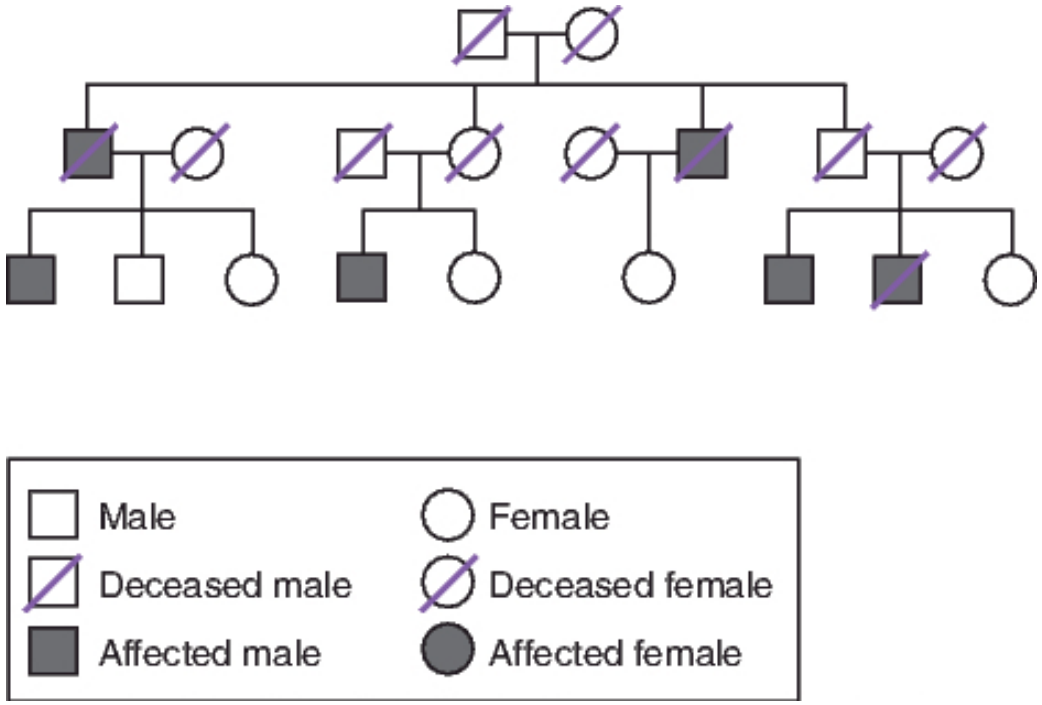
- **Wildtype** (**WT**)
  The prevailing characteristic, or gene, of a population
  May also be referred to as “normal” or “negative”
- **Mutant** (**MUT**)
  The variant form of a gene, not WT
  May also be referred to as “abnormal” or “positive”
- **Homozygous** (**HOM**)
  Possessing two of the same allele (*e.g.*, tt and TT)
- **Heterozygous** (**HET**)
  Possessing different alleles (*e.g.*, Tt)



Reading a Pedigree

A **pedigree** is a diagram of the inheritance pattern of a specific genotype/phenotype of a family.

The construction and analysis of a pedigree is used to determine the **pattern of inheritance**, also called **transmission pattern**, of a disease.



Patterns of Inheritance

Mendel's Laws of Inheritance	
Law of Dominance and Uniformity	Some alleles are dominant while others are recessive. An organism with at least one dominant allele will display the effect of the dominant allele.
Law of Segregation	During gamete formation, allele pairs for each gene segregate from each other so that each gamete carries only one allele for each gene.
Law of Independent Assortment	Genes of different traits can segregate independently during the formation of gametes.

Patterns of Inheritance

Mendelian inheritance patterns refer to those where genes express predictably among offspring based on Gregor Mendel's original laws of inheritance.

- Is the gene locus on an autosome or sex chromosome?
- Is the gene dominant or recessive?

Non-mendelian inheritance patterns are those which violate one or more of Mendel's laws (more complicated than a Punnett Square).

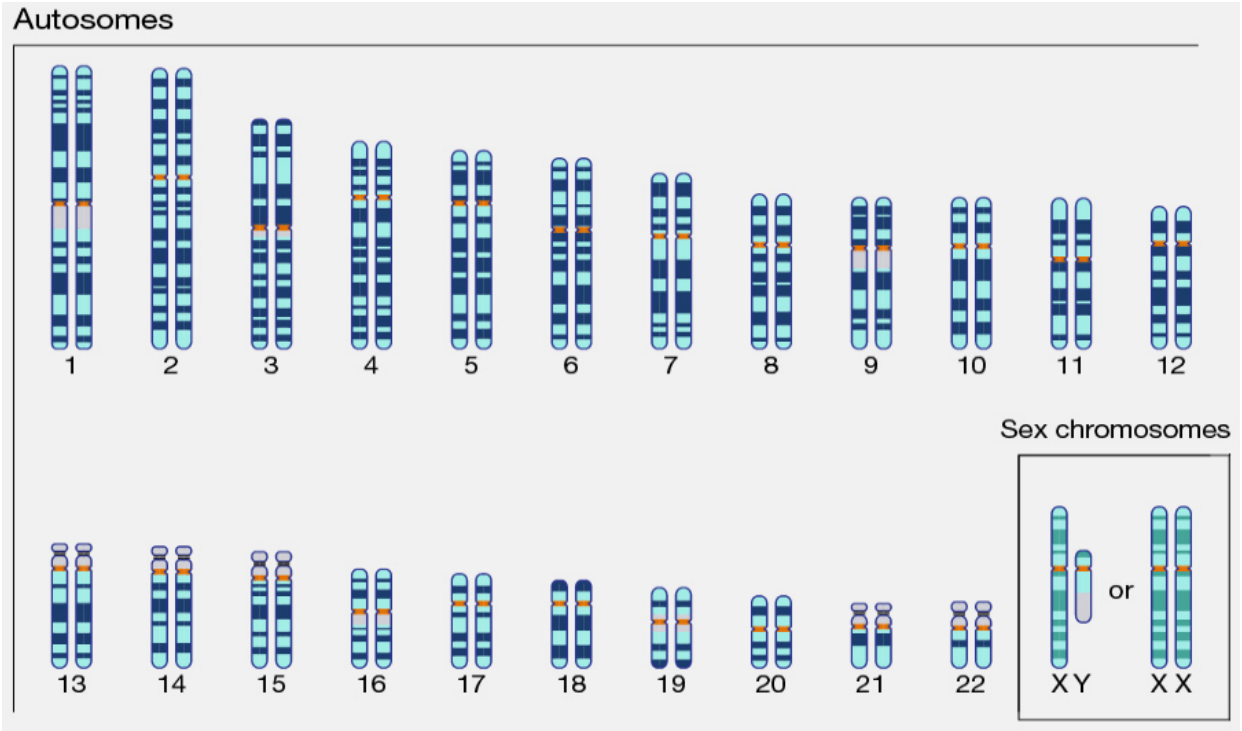
Mendelian	Non-Mendelian
Autosomal dominant	Codominant
Autosomal recessive	Mosaic
X-linked dominant	Genomic imprinting
X-linked recessive	Trinucleotide repeat
	Mitochondrial

Mendelian Patterns of Inheritance

For **autosomal disorders**, the abnormal allele is carried on one of the 22 **autosomes**.

Autosomes are distributed equally among the sexes, so female (46,XX) and male (46,XY) patients are affected at the same rate.

Expression of the disorder is dependent on whether the abnormal allele is **dominant** or **recessive**.

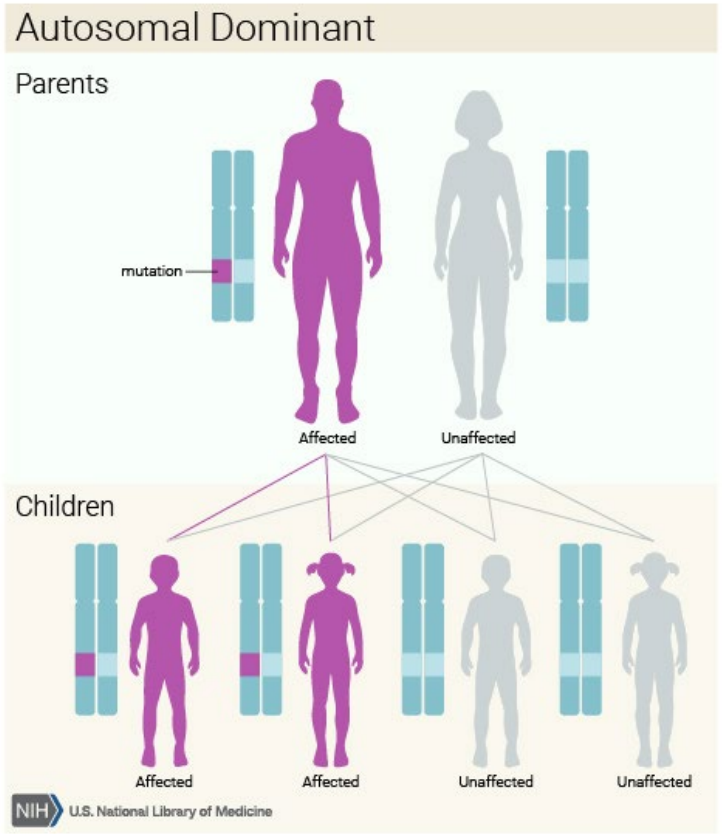


Mendelian Patterns of Inheritance

For **autosomal dominant** disorders, a single abnormal allele is enough for a patient to be affected.

- WT-WT (wildtype) = unaffected
- MUT-WT (heterozygous) = affected
- MUT-MUT (homozygous) = affected

Example: Huntington's Disease



Mendelian Patterns of Inheritance

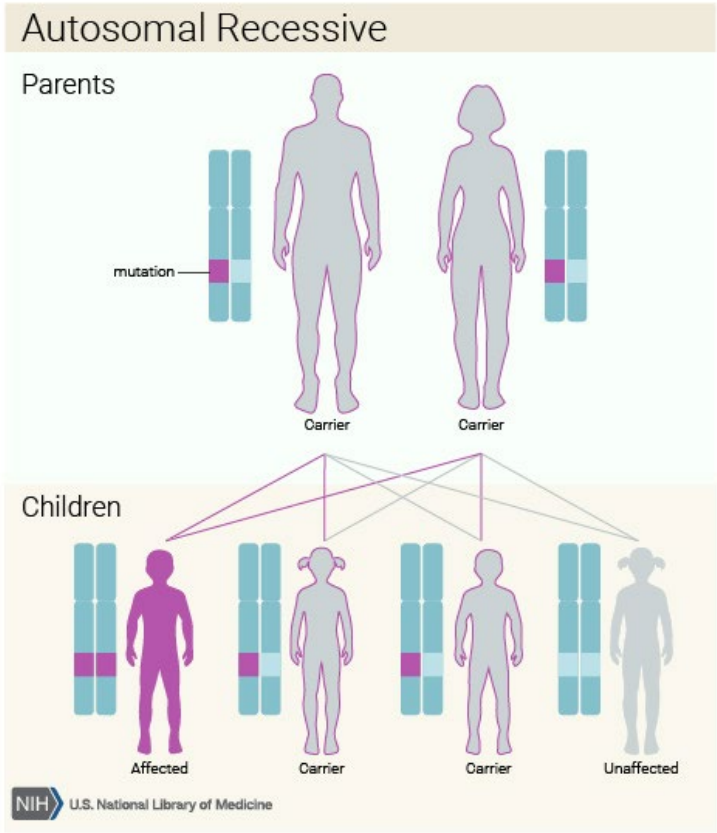
For **autosomal recessive** disorders, both alleles must be abnormal for a patient to be affected

- WT-WT (wildtype) = unaffected
- MUT-WT (heterozygous) = unaffected
- MUT-MUT (homozygous) = affected

Carrier is the term used to refer to people who have only one abnormal recessive allele (heterozygotes).

- Unaffected by disease
- Still have the ability to pass disease to their offspring

Example: Cystic Fibrosis

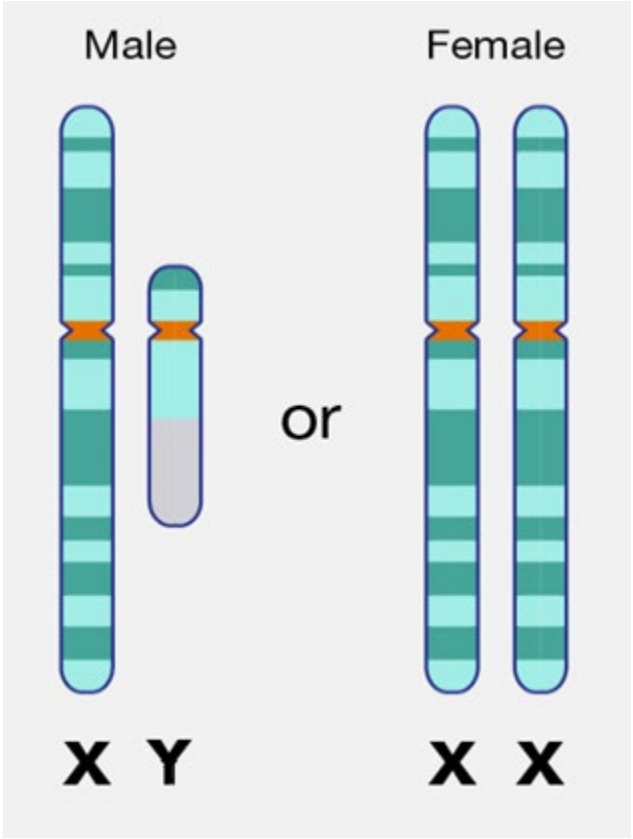


Mendelian Patterns of Inheritance

For **sex-linked disorders**, the abnormal gene is carried on one of the sex chromosomes (generally the X chromosome).

Because X chromosomes are distributed unequally among the sexes, **X-linked disorders** are expressed unequally as well.

- Males (46,XY) are **hemizygous** for the X chromosome. Regardless of whether the disease is recessive or dominant, all males that inherit an abnormal allele will be affected.
- Females (46,XX) have two X chromosomes. Disease expression is more variable dependent on allelic configuration and recessive/dominant nature of the disorder.

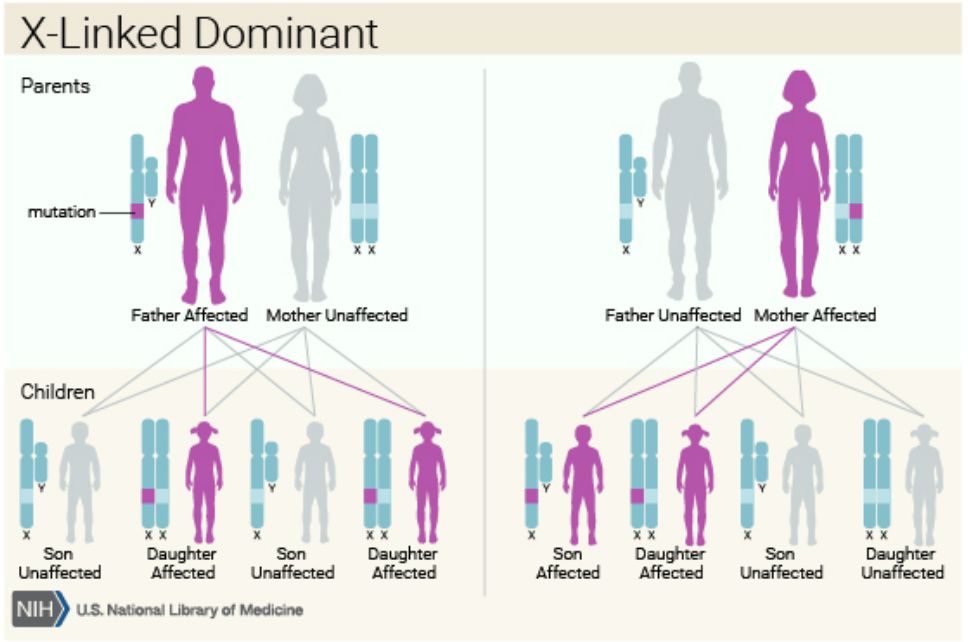


Mendelian Patterns of Inheritance

For **X-Linked dominant** disorders, a single abnormal X chromosome is enough for a patient to be affected.

X inactivation, a process where female (46,XX) cells randomly “turn off” the gene activity of one of their X chromosomes, alters disease expression in heterozygotic females when compared to homozygotic females and hemizygotic males.

Example: Fragile X Syndrome

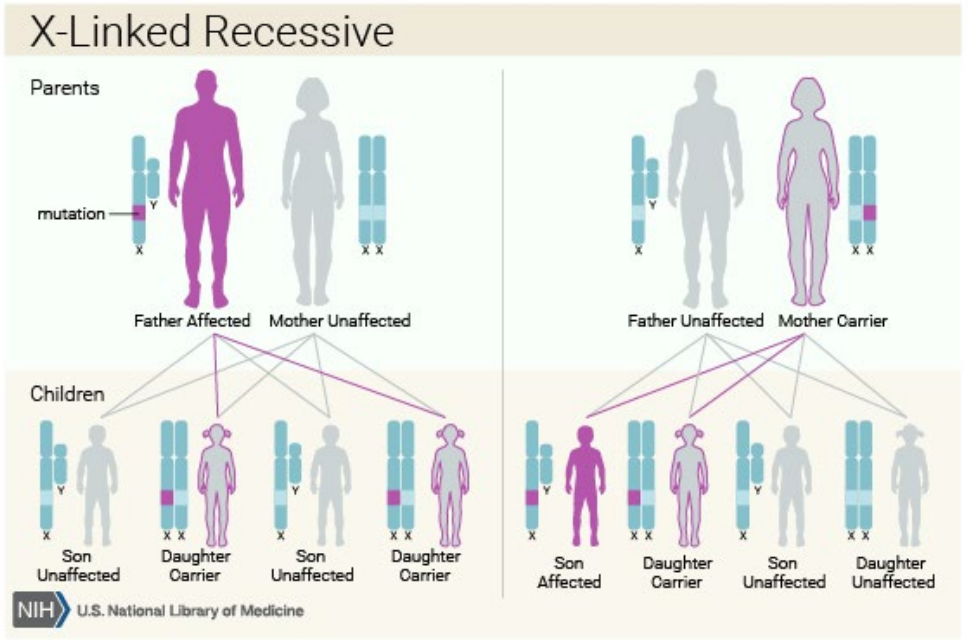


Mendelian Patterns of Inheritance

For **X-linked recessive** disorders, all X chromosomes must be abnormal for a patient to be affected.

- For females (46, XX), both alleles must be abnormal for patient to be affected.
- Males (46, XY) are **hemizygous** for the X chromosome, so a single abnormal allele is enough to present disease.

Example: Color blindness

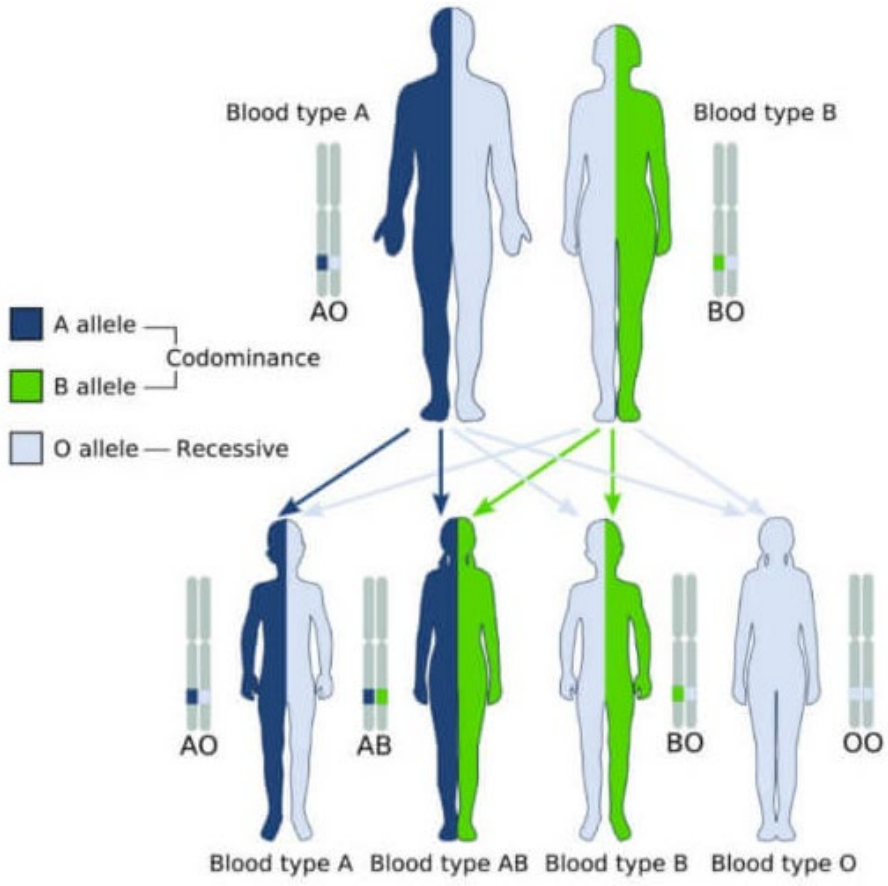


Non-Mendelian Patterns of Inheritance

Codominance is when the phenotype of two different alleles are both expressed in heterozygotes.

No individual version of a gene is dominant over another.

Example: ABO blood group system

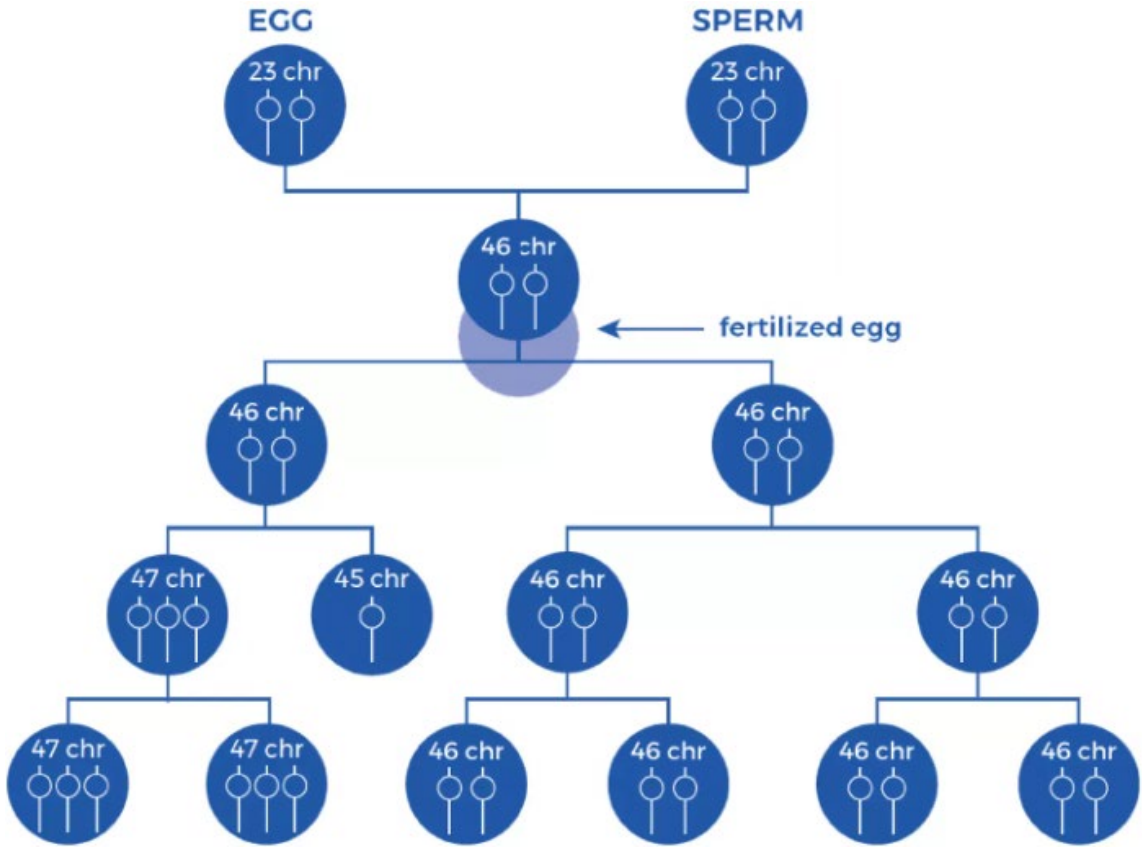


Non-Mendelian Patterns of Inheritance

Mosaicism is the presence of two or more cell lineages that differ genetically but are derived from a single zygote. Mutation is not present ubiquitously, with only some of a patient's cells exhibiting mutation.

- Caused by mitotic nondisjunction early in embryogenesis.

Example: ~2% of Down Syndrome (Trisomy 21) cases are mosaic



Non-Mendelian Patterns of Inheritance

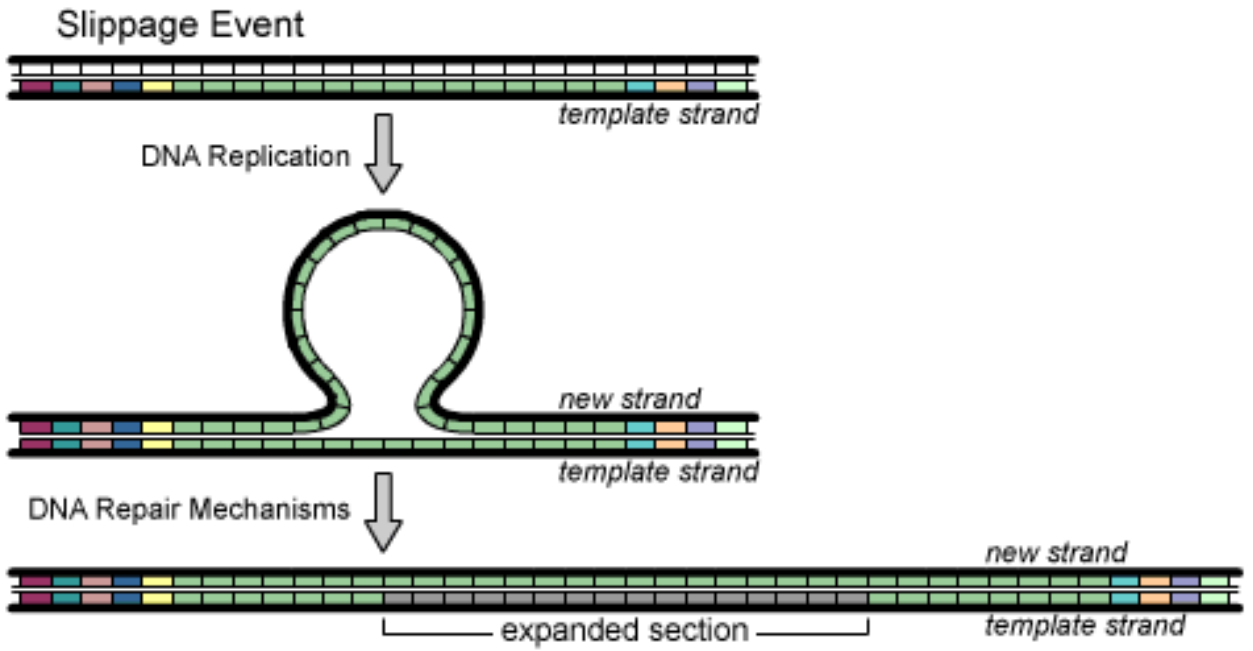
Trinucleotide repeat disorders are those where 3-base pair STRs expand in length during DNA replication and meiosis.

- Repeat expansions alter/silence gene expression.

Each subsequent generation has an increased risk of further repeat expansion and more severe disease presentation, a phenomenon called **anticipation**.

- "Slippage" phenomenon

Example: Fragile X Syndrome, Huntington's Disease



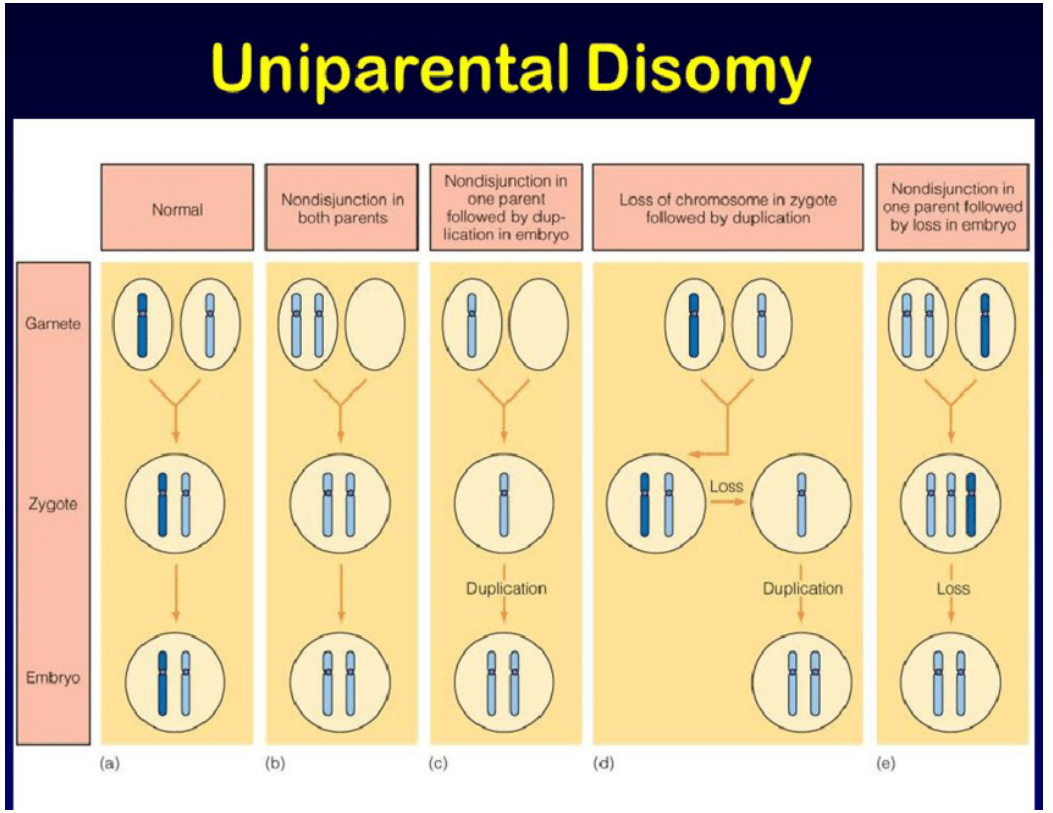
Non-Mendelian Patterns of Inheritance

Genomic imprinting refers to transcriptional silencing through histone/DNA modification that occurs during gamete production.

- Gene silencing differs for DNA in eggs and DNA in sperm.
- Disorders occurs when either the allele from the egg or the allele from sperm is lost.
- **Uniparental disomy** is when both alleles are inherited from the same parent.

Disease expression differs depending on maternal vs. paternal transmission.

Example: Prader-Willi vs. Angelman Syndrome



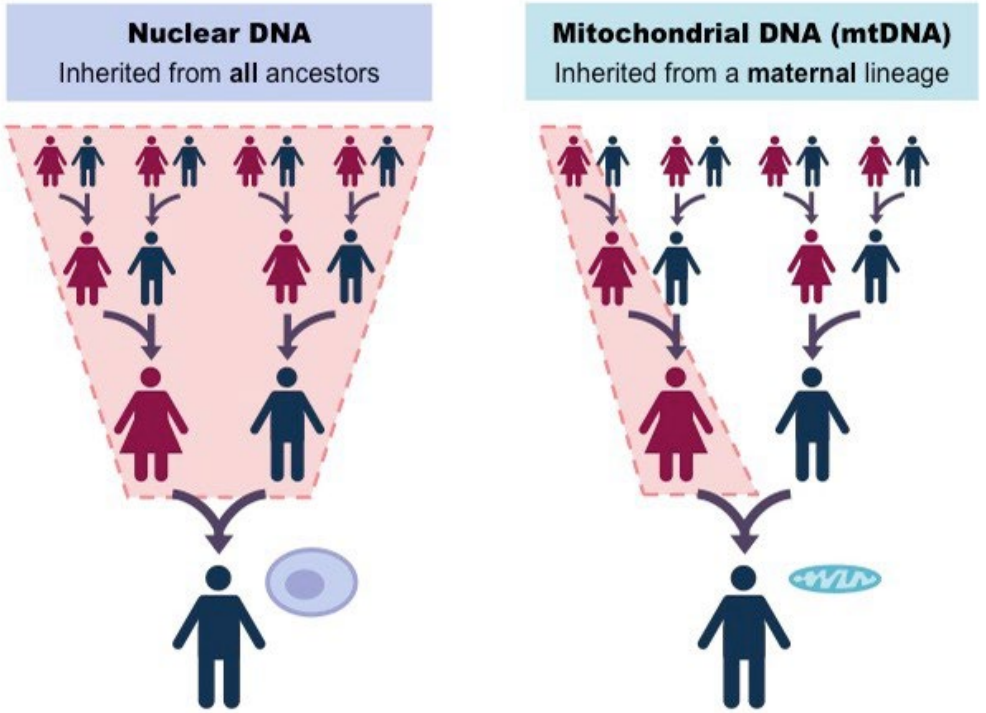
Non-Mendelian Patterns of Inheritance

Mitochondrial disorders are those arising from mutations in the mitochondrial genome.

- Mitochondria, the powerhouse of the cell, have their own separate genome from that found in the nucleus.

Mitochondrial DNA (mtDNA) is only passed from mother to offspring, so mitochondrial disorders have strict maternal inheritance.

Example: Myoclonic Epilepsy with Ragged-Red Fibers (MERRF)



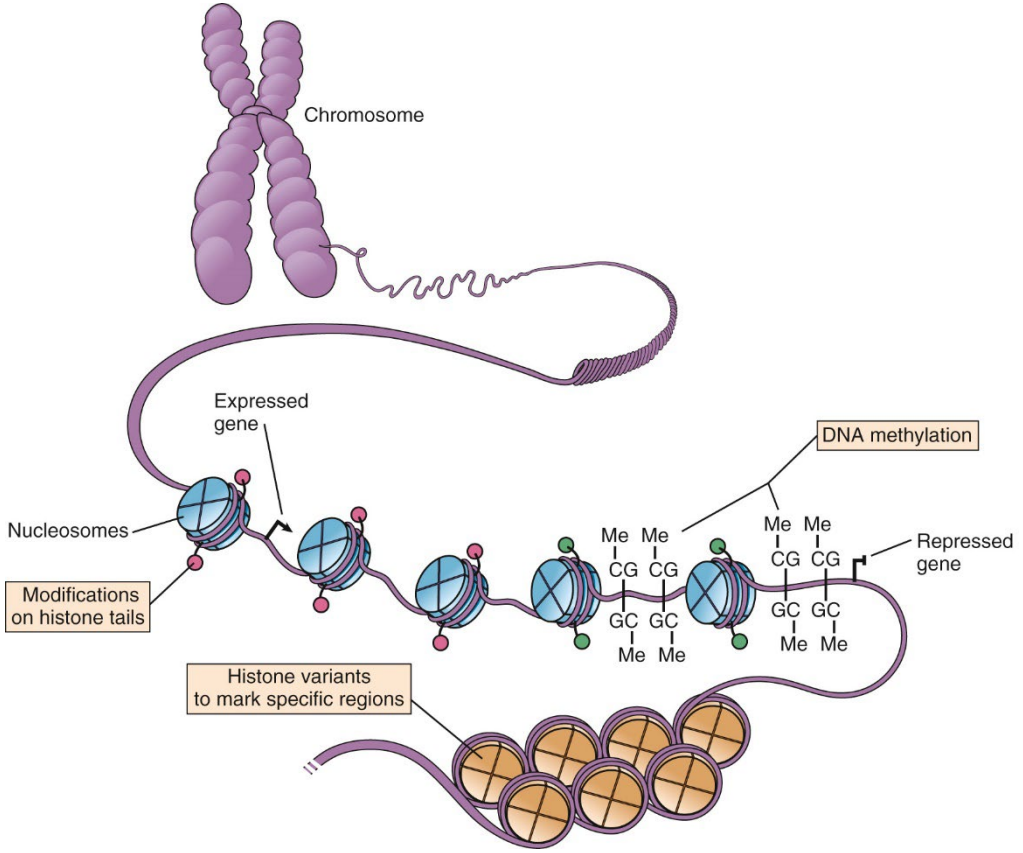
Epigenetic Factors

Not all phenotypic variation is due to sequence-based changes to the genome.

Epigenetic factors are those which alter gene expression through up- and down-regulation of gene transcription.

There are three primary epigenetic mechanisms:

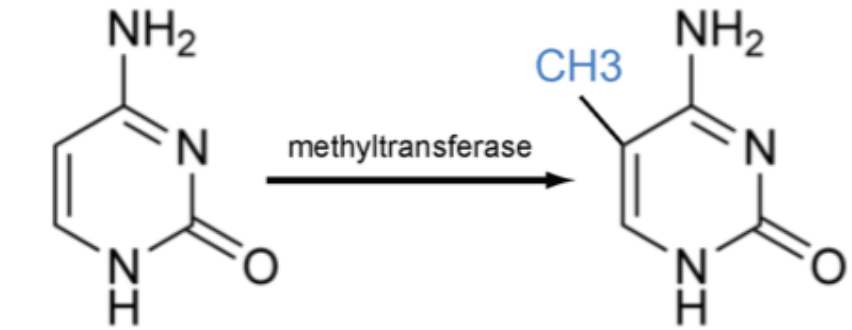
- Histone modification
- Genomic Imprinting
- **DNA methylation**



DNA Methylation

DNA methylation occurs when a methyl group is added to cytosine bases.

- DNA methyltransferases recognize cytosine bases in DNA followed by guanine bases (CpG).
- **CpG islands**, areas of the genome containing lots of CpGs, are frequently found in the first exons and promoter regions of genes.
- Hypermethylation of CpG dinucleotides can reduce transcriptional activity of a gene or silence it altogether.



Unmethylated cytosine

Methylated cytosine

...GGAGGAGCGCGCGGCGGCGGCCAGAGA

AAAGCCGCAGCGGCGCGCGCGCACCCGGA

CAGCCGGCGGAGGCGGG...

Example CpG island sequence showing a higher-than-expected occurrence of CpG

Questions?

EMAIL: POHLJ@CCF.ORG

WORK PHONE: 216 308-0801

This concludes the presentation.

CLEVELAND CLINIC CENTER FOR PATHOLOGY EDUCATION