Overview

- Basic Terminology
- Introduction to the Human Genome
- Genetic Variation
- Genetic Tests
- Whole Genome Sequencing
- Variant Interpretation

Basic Terminology

- Genome Biology
- Genomic Medicine
- Genetic Reports
- Variant Classification (SNV)
- Variant Classification (CNV)

Basic Terminology: Genome Biology

- **Deoxyribonucleic acid (DNA)**: hereditary material in all organisms, located in the nucleus and mitochondria, and identical in every cell of an individual
- **Nucleotide**: building block of DNA, 4 possible nucleotides: A, T, C, G
- **Gene**: a functional unit of genetic information
- **Allele**: different forms or versions of a gene
- **Chromosome**: natural state of condensed DNA in the cell nucleus, packaged into 46 chromosomes (22 pairs of autosomes and 2 sex chromosomes)
- **Genome**: the entire set of genetic information, including both coding and noncoding regions
- **Exome**: the subset (1-2%) of the genome that provides instructions for building proteins
- **Genotype**: the nucleotide or string of nucleotides at a specific location in an individual's genome
- **Phenotype**: a patient's clinical presentation or features

Basic Terminology: Genomic Medicine

- **Genomic medicine**: clinical management informed by genomic information
- **Genetic disease**: a disease caused by a DNA change(s); also called a genetic condition or genetic disorder
- **Rapid whole genome sequencing**: determining an individual's genetic sequence in a rapid timeframe (3-7 days)
- **Variant**: a change in the DNA sequence compared to the "normal" reference genome; may be pathogenic (disease causing) or benign; may be a single nucleotide substitution (missense, nonsense), deletion or duplication of one or more nucleotides (frameshift), deletion or duplication of gene regions, genes, or genomic segments, or rearrangement of genomic regions
- **Homozygous**: refers to identical alleles at a specific DNA position or locus
- **Heterozygous**: refers to different alleles at a specific DNA position or locus
- **Proband**: the affected individual (patient)
- **Trio**: the affected individual (patient) and his/her parents

Basic Terminology: Genetic Reports

- **Primary findings**: genetic findings related to the patient's clinical presentation
- **Incidental findings**: genetic findings not related to the patient's clinical presentation
- **Autosomal**: relating to a chromosome that is not a sex chromosome (chromosomes 1-22)
- **X-linked**: relating to the X chromosome
- **Mitochondrial**: relating to the mitochondria, powerhouses of the cell that contain their own DNA
- **Dominant**: a genetic condition expressed when a defect in **just one copy** of a gene is sufficient to cause disease
- **Recessive**: a genetic condition expressed when a defect in **both copies** of a gene is necessary to cause disease
- **De novo**: variant is only detected in the affected individual, not in the parents

Terminology: Variant Classification (SNV)

- **Pathogenic**
- **Likely pathogenic** (>90% certainty)
- **Variant of uncertain clinical significance (VUS)**
- **Likely benign** (>90% certainty)
- **Benign**

Terminology: Variant Classification (CNV)

- Pathogenic
- VUS, likely pathogenic
- Variant of uncertain significance (VUS)
- VUS, likely benign
- Benign

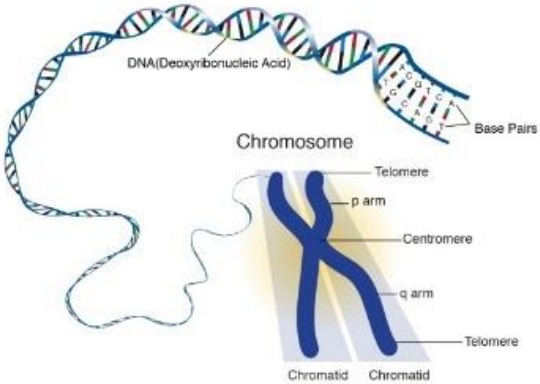
Introduction to the Human Genome

- DNA Structure
- Genome Structure and Organization
- Gene Structure and Organization
- Central Dogma of Molecular Biology
- Gene Expression (Transcription and Translation)
- Epigenetics and Epigenomics

The Human Genome

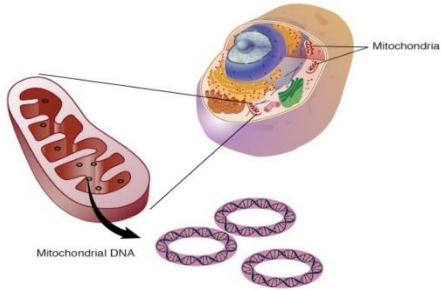
- The human **genome** consists of large amounts of DNA that specifies all aspects of embryogenesis, development, growth, and metabolism
- Every nucleated cell in the body carries its own copy of the human genome
- **Genes** are encoded in DNA and the genome is organized into **chromosomes**
- The nuclear genome consists of 46 chromosomes:
    - 22 pairs of autosomes (numbered 1 to 22, in order of size)
    - 1 pair of sex chromosomes (XX or XY)
- A small but important part of the human genome resides in the mitochondria

DNA Structure



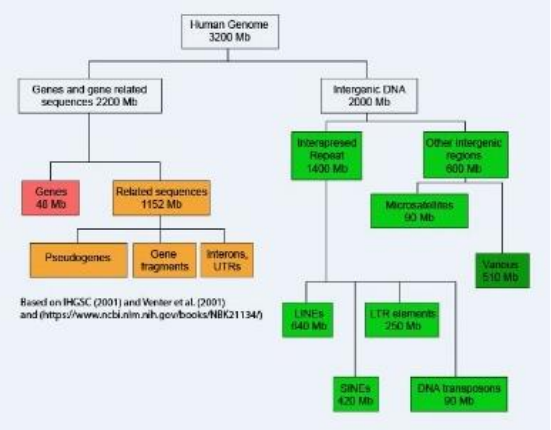
- DNA is a macromolecule composed of nucleotides: A, T, C, G
- DNA exists in the form of a double helix, in which the two strands run in opposite directions and are held together by hydrogen bonds
- The double-stranded complementary structure of DNA allows for the precise replication of DNA and the accurate repair of damaged DNA

Genome Structure



- The nuclear genome is packaged as **chromatin** in the nucleus
- Chromatin consists of nucleosomes (DNA bound to histone proteins)
- The mitochondrial genome resides in the mitochondria
- Mitochondria and mitochondrial DNA exhibit exclusively maternal inheritance

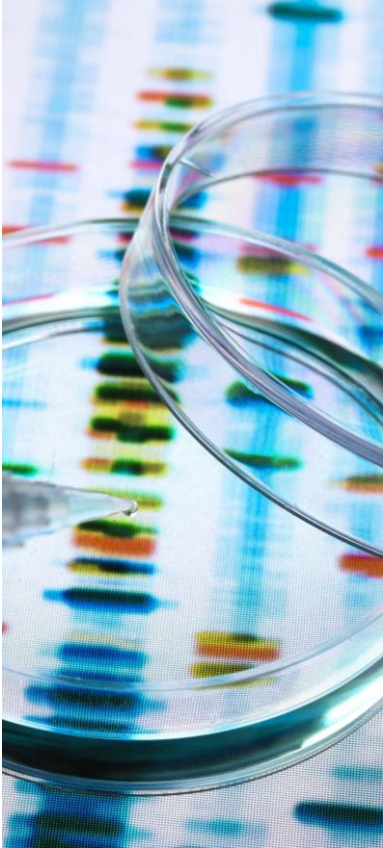
The Human Genome Project



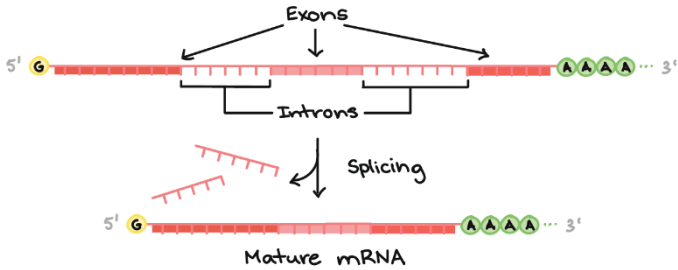
- The Human Genome Project was an international effort to determine the sequence of the 3 billion base pairs of DNA within the human genome
- The consensus sequence was reported in 2003 as a **"reference" sequence**
- The reference sequence is maintained in publicly accessible databases and is used as a basis for comparison with sequences of individual genomes

Organization of the Human Genome

- **Exons** (regions of genes coding for proteins or functional RNAs) make up **1.5%** of the genome
- **Introns** (regions of genes not coding for proteins or functional RNAs) and regulatory sequences account for **24%** of the genome
- Much of the genome consists of **repetitive DNA sequences**, including tandem repeats (STRs or microsatellites) and interspersed repeats (Alu, LINE, SINE)



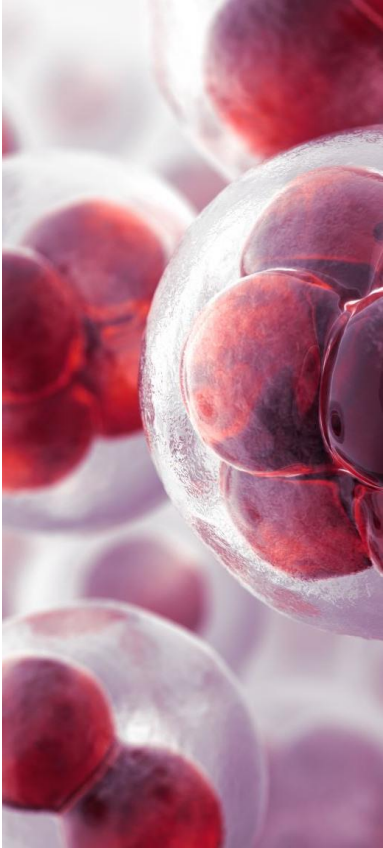
Gene Organization and Structure



- A **gene** is a sequence of DNA that specifies production of a functional product (protein or noncoding RNA)
- A gene includes the coding sequences (exons) and adjacent sequences required for proper gene expression (introns, promoters, regulatory elements)

Central Dogma of Molecular Biology

- The **Central Dogma of Molecular Biology** refers to the flow of information from DNA to RNA to protein
- Genetic information is stored as **DNA** in the nucleus
- **RNA** is synthesized from the DNA template by RNA polymerase through a process known as **transcription**
- The resulting messenger RNA (mRNA) is then transported from the nucleus to the cytoplasm where the mRNA is decoded by the ribosome into a **protein** product in a process known as **translation**



Transcription: DNA to RNA

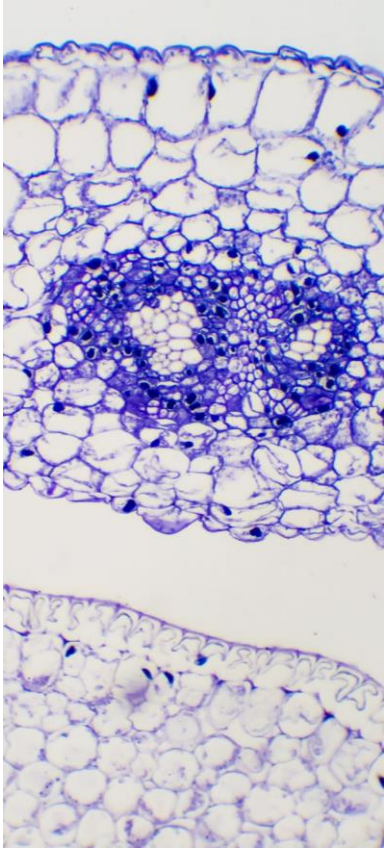
- **Transcription** of protein-coding genes is initiated at the transcriptional start site by RNA polymerase II
- Synthesis of RNA transcripts proceeds in a 5’ to 3’ direction
- The **template strand** of DNA is actually read in a 3’ to 5’ direction (antisense strand) while the **non-template strand** of DNA corresponds both in polarity and sequence (sense strand) to the synthesized RNA
- The primary RNA transcript is processed by addition of a “cap” to the 5’ end and a polyA tail to the 3’ end

RNA Splicing

- After transcription, introns are removed and exons are joined together in a process called **RNA splicing**
- Splicing is guided by specific sequences in the primary RNA transcript at the junctions of exons and introns (**splice sites**)
- The primary RNA transcript can follow multiple **alternative splicing** pathways, leading to the synthesis of different mRNAs and protein products

Translation: RNA to protein

- After transport from the nucleus to the cytoplasm, mRNA is translated into protein by the ribosome
- Transfer RNA (tRNAs) bring the correct amino acids into position along the mRNA transcript
- Ribosomes, composed of ribosomal RNAs (rRNAs) and proteins, catalyze protein synthesis



Genetic Code

- Translation of genetic information is determined by the **genetic code**, which relates specific amino acids to combinations of three adjacent bases (codons)
- Because there are only 20 amino acids and 64 (4^3) possible codons, most amino acids are specified by more than one codon
- Translation is always initiated by a methionine (Met or M) and translation ends when a stop (nonsense) codon is encountered

Epigenetics and Epigenomics

- **Epigenetics** refers to the reversible, heritable changes in the chromatin landscape (rather than changes to the genome sequence itself) that can influence gene expression
- Complex epigenetic states can be established and maintained by a variety of mechanisms:
  - DNA methylation
  - Histone modification
  - Histone variants

Genomic Imprinting

- Imprinting is a normal process involving the introduction of epigenetic marks in the germline of one parent, but not the other, at specific locations in the genome
- Imprinting leads to monoallelic expression (expression from one allele) of a gene or multiple genes within the imprinted region
- Defects in genomic imprinting can result in human disease
    - Example: Prader-Willi Syndrome & Angelman Syndrome

Genetic Variation

- **Types of Genetic Variants**
  - SNVs
  - Small indels
  - CNVs
  - SVs
  - Aneuploidy
- **Patterns of Single-Gene Inheritance**
  - Dominant and Recessive Traits
  - Penetrance and Expressivity
  - Mosaicism
  - Uniparental Disomy

Genetic Variation

- The prevailing version of a gene (allele) is the **wild-type** (WT) allele
- Other versions of a gene are **variant** alleles
- The frequency of different variants can vary widely in different populations
- Changes to DNA can be classified by their size and nature:
    - Single nucleotide variants (SNVs)
    - Small insertions and deletions (indels)
    - Copy number variants (CNVs)
    - Structural variants (SVs)
    - Aneuploidy
- Depending on the location, nature, and size of the alteration, the functional consequence of a DNA change can be **benign** or **deleterious (pathogenic)**

Types and Consequences of Genetic Variants

Type	Variant	Consequence/Effect
Single nucleotide variant	Missense (nonsynonymous) Nonsense (stop gained) Splice site	Change in protein sequence Termination and NMD Altered splicing
Small insertions or deletions	Frameshift In-frame insertion or deletion Repeat expansion	Termination and NMD Change in protein length Change in protein length
Copy number variant	Gene dup or del Chromosomal region dup or del	Altered gene dosage Altered gene(s) dosage
Structural variant	Inversion Rearrangement (translocation)	Altered gene expression Altered gene expression/dosage
Aneuploidy	Trisomy, monosomy	Altered chromosome dosage

Single Nucleotide Variants (SNVs)

- SNVs are substitutions of individual nucleotides in the genome
    - Examples: CHD7 (CHARGE Syndrome), KMT2D (Kabuki Syndrome)
- SNVs can have multiple effects:
    - Change the amino acid sequence (missense)
    - Introduce a stop codon (nonsense, stop gained)
    - Alter a splice site

Small Insertions and Deletions (indels)

- Indels refer to the insertion or deletion of a single or multiple base pairs
- Indels can have multiple effects:
  - Multiple of 3 nucleotides: inframe insertion or deletion, resulting in a change of protein length
  - Non-multiple of 3 nucleotides: frameshift leading to a stop codon and early termination
- Expansion of trinucleotide repeats beyond the normal range can result in repeat expansion disorders
  - Example: Huntington disease (CAG repeat)

Copy Number Variants (CNVs)

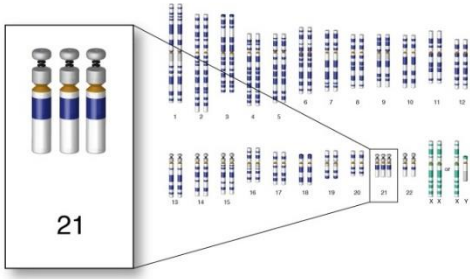
- Copy number variants (CNVs) consist of variation in the number of copies of large segments of the genome (range: 1000 bp to hundreds of kb)
- Deletion: less than 2 copies of a genomic region
    - Example: 22q11.21 del (DiGeorge syndrome)
- Duplication: more than 2 copies of a genomic region
- Many CNVs are found in regions with repeated homologous sequences called segmental duplications (segdups)

Structural Variants (SVs)

- Structural variants (SVs) refer to variation in the structure, location, or number of chromosomal regions
- SVs include CNVs, inversions, and rearrangements (translocations)
- Inversions occur when a chromosome undergoes two breaks and is reconstituted with the segment inverted, usually do not result in a change in copy number, and can occur within one chromosome arm (paracentric) or include the centromere (pericentric)
- Rearrangements result from chromosome breakage followed by reconstitution in an abnormal combination and can result in normal copy number (balanced) or additional or missing genomic material (unbalanced)
- Translocations involve the exchange of chromosome segments between two chromosomes and can result in reciprocal exchange of chromosome segments (reciprocal translocation), fusion of two long arms of chromosomes (Robertsonian translocation), or additional or missing genomic material (unbalanced translocation)

Aneuploidy

- Aneuploidy is any variation in chromosome number
- Monosomy: 1 copy of a particular chromosome
    - Example: Monosomy X (Turner Syndrome)
- Trisomy: 3 copies of a particular chromosome
    - Examples: Trisomy 13, Trisomy 18, Trisomy 21 (Down Syndrome)



Origin of Genetic Variation

- Variants arise in the context of normal cellular processes
- SNVs and indels can originate from errors in DNA replication or DNA repair
- CNVs and SVs can arise during homologous recombination (HR) or non-homologous end joining (NHEJ)
- Aneuploidy can arise from chromosome missegregation during cell division
- All inherited genomic variation had to originate as a new (de novo) change occurring in germ cells (germline)
- Variation that occurs in developed cells of the body (somatic) cannot be transmitted to offspring, but can cause cancer if the variant imparts a growth advantage to the cell

Patterns of Single-Gene Inheritance

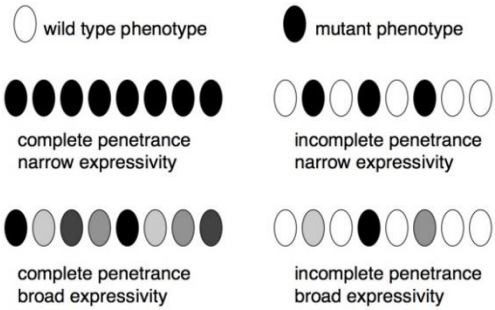
- Mendelian (single-gene) disorders are determined primarily by alleles at a single genomic locus
- Classic Mendelian inheritance patterns:
  - Autosomal dominant (AD)
  - Autosomal recessive (AR)
  - X-linked dominant (XLD)
  - X-linked recessive (XLR)
- Mendelian disorders occur on average in fixed and predictable proportions among offspring of specific types of matings
- Genetic variants often exhibit pleiotropy (multiple diverse phenotypic effects in multiple organ systems from a single genetic variant)

Dominant and Recessive Traits

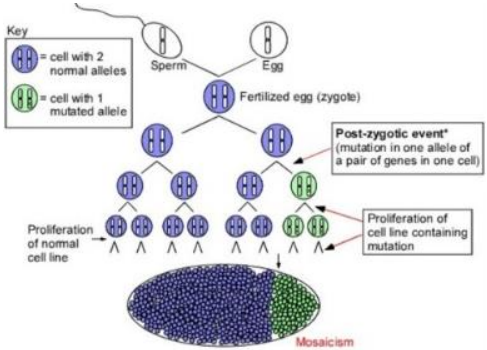
- Autosomal dominant
  - Expressed in heterozygotes, homozygotes, or compound heterozygotes
- Autosomal recessive
  - Expressed **only** in homozygotes or compound heterozygotes
- X-linked dominant
  - Expressed in heterozygotes (females) and hemizygotes (males)
- X-linked recessive
  - Expressed **only** in hemizygotes (males)

Penetrance and Expressivity

- Some genetic conditions are always full expressed at birth, but other disorders are not expressed at all or may vary substantially in symptoms, severity, or age of onset
- Penetrance: probability that a variant will have any phenotypic expression
- Expressivity: severity of expression of a phenotype



Mosaicism



- Mosaicism refers to the presence of two or more different chromosome complements in an individual or tissue
- Variants that occur after conception in a single cell can give rise to clones of cells genetically different from the original zygote
- The population of cells that harbor a variant might be present in some tissues (somatic mosaicism) or restricted to the gamete lineage (germline mosaicism)

Uniparental Disomy (UPD)

- Uniparental disomy (UPD) occurs when both copies of a chromosome derive from the same parent
  - Isodisomy: derived from identical sister chromatids
  - Heterodisomy: derived from different homologues from one parent
- Uniparental disomy arises from trisomy rescue due to chromosome nondisjunction to restore a disomic state (2 chromosomes)

Mitochondrial Disorders

- Mitochondrial DNA (mtDNA) and mitochondrial disorders manifest strictly maternal inheritance because only maternal mitochondria are transmitted to offspring
- During cell division, mitochondria are randomly segregated into daughter cells, resulting in a mixture (heteroplasmy) or a pure population (homoplasmy) of mitochondrial DNA
- Phenotypic expression of a mtDNA variant depends on plasmy

Types of Genetic Tests

- Karyotype
- Fluorescence *in situ* hybridization (FISH)
- Microarray
- Single gene or gene panel testing
- Whole exome sequencing (WES)
- Whole genome sequencing (WGS)

Karyotype

- **Purpose**: to visualize and detect aneuploidy (loss or gain of chromosomes) or structural variants (deletions, duplications, or translocations larger than 5 MB)
- **Advantages**: can detect large chromosomal abnormalities across the genome, especially unbalanced rearrangements
- **Disadvantages**: cannot detect small genetic variants, including SNVs, indels, and deletions or duplications smaller than 5 MB

Fluorescence *in situ* Hybridization (FISH)

- **Purpose**: to detect structural variants (deletions, duplications, inversions, and translocations) at a known genomic locus using fluorescent probes
- **Advantages**: can detect large chromosomal abnormalities at a specific locus
- **Disadvantages**: cannot detect small genetic variants, DNA region of interest must be known, not genome-wide

Microarray (array CGH or SNP array)

- **Purpose**: to detect copy number changes (deletions or duplications) across the genome using probe-based arrays on a chip; often confirmed by FISH
- **Advantages**: can detect copy number changes (small to large) across the genome
- **Disadvantages**: cannot detect small genetic variants or unbalanced rearrangements

Single Gene or Gene Panel Testing

- **Purpose**: to detect genetic variants (SNVs, indels, deletions, duplications) in a single gene or small number of genes (gene panel) associated with a suspected genetic condition
- **Advantages**: can detect multiple types of genetic variants in specific genes, high coverage for genes of interest, cost-effective
- **Disadvantages**: limited to specific genes of interest, not genome-wide

Whole Exome Sequencing (WES)

- **Purpose**: to detect genetic variants (SNVs, indels, deletions, duplications) across the exome (protein-coding subset of the genome)
- **Advantages**: can detect multiple types of genetic variants in the exome; high coverage across the exome (>100x)
- **Disadvantages**: cannot detect large chromosomal abnormalities (aneuploidy, rearrangements, large deletions or duplications) or genetic variants in non-coding regions (98% of the genome); little or no coverage of regions not covered by exome capture probes; high cost

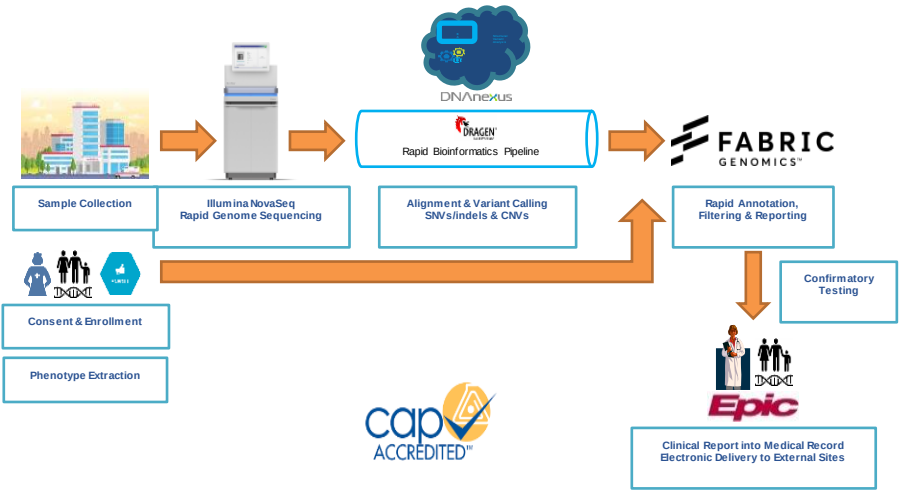
Whole Genome Sequencing (WGS)

- **Purpose**: to detect genetic variants (SNVs, indels, deletions, duplications, rearrangements, aneuploidy) across the genome (including introns and noncoding regions)
- **Advantages**: can detect multiple types of genetic variants in the genome (SNVs, indels, deletions, duplications, rearrangements, aneuploidy); high coverage across the genome (30-40x); more uniform coverage compared to whole exome sequencing
- **Disadvantages**: cannot detect balanced rearrangements; large amounts of data; difficulty in interpretation of variants in noncoding regions; high cost

Comparison of Genetic Tests

	Detects SNV/indel?	Detects CNV?	Coverage	Resolution	Turn-Around Time
Karyotype		Yes	Genome	Low	3-10 days
FISH		Yes	Targeted	High	3-10 days
Microarray		Yes	Genome	Medium	10-14 days
Gene panel	Yes	Yes	Targeted	High	4-6 weeks
Whole exome sequencing (WES)	Yes	Yes	Exome	High	8-12 weeks
Whole genome sequencing (WGS)	Yes	Yes	Genome	High	6 weeks; 3-14 days (RCIGM)

Overview of RCIGM Rapid WGS Workflow



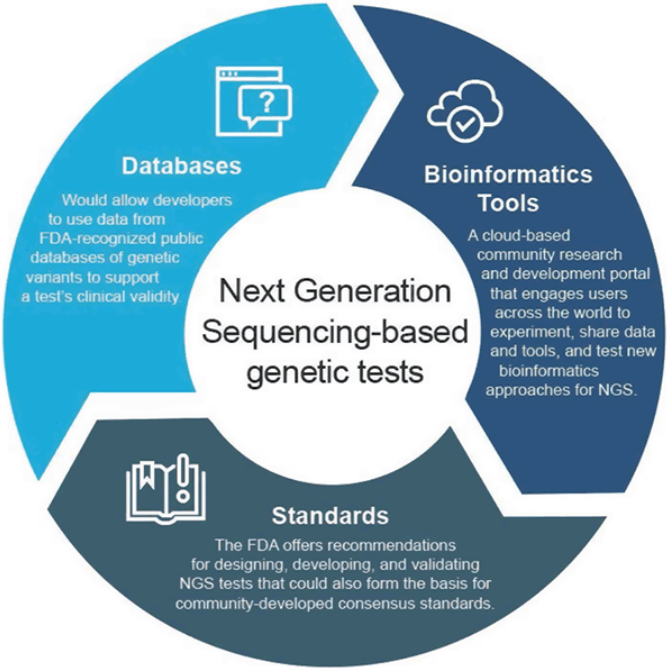
Overview of Whole Genome Sequencing (WGS)

- DNA Isolation
- Library Preparation
- Sequencing
- Read Alignment
- Variant Calling
- Variant Annotation

DNA Isolation

- DNA is typically extracted from blood, saliva, or solid tissue
- This process involves:
  - Lysis of cells
  - Purification of DNA from other cellular material
  - Precipitation and elution of DNA in a buffer
  - Quantification of DNA

Next-Generation Sequencing Overview



Library Preparation

- Library preparation involves the generation of a collection of DNA fragments for sequencing
- This process involves:
  - Fragmentation of genomic DNA by sonication or enzymatic activity
  - Ligation of specialized adapters to both fragment ends
- Indexes can be added to genomic DNA fragments to allow for multiplexing, the pooling and sequencing of multiple samples together in a single flow cell

Cluster Generation

- Once the sequencing library is loaded onto a flow cell, the fragments are hybridized to complementary oligos bound to the flow cell surface
- Each fragment is then amplified into distinct, clonal clusters through bridge amplification
- When cluster generation is complete, the templates are ready for sequencing

Sequencing

- Sequencing by synthesis (SBS) technology uses a reversible terminator-based method that detects single bases as they are incorporated into DNA strands
- As each base is incorporated, the flow cell is imaged and the emission from each cluster is recorded
- The emission wavelength and intensity are used to identify the base
- Cycles are repeated to create “reads”
- The output is a fastq file, which contains raw (unaligned) sequencing reads

Read Alignment and Variant Calling

- The newly identified sequence reads are aligned to a reference genome using bioinformatics software
- The output is a bam file, which contains aligned reads
- Computer algorithms compare the aligned sequencing reads to the reference genome to determine the presence or absence of SNVs and indels
- The output is a vcf file, which contains a list of genetic variants

Variant Annotation

- Genetic variants are annotated using information from genomic databases
  - Gene
  - Transcript
  - Genomic position
  - cDNA and protein change
  - Zygosity
  - Allele frequency in population databases (ExAC, gnomAD)
  - Previous reports in variant databases (ClinVar, HGMD)
  - Computational scores (in silico predictions)

Overview of Variant Interpretation

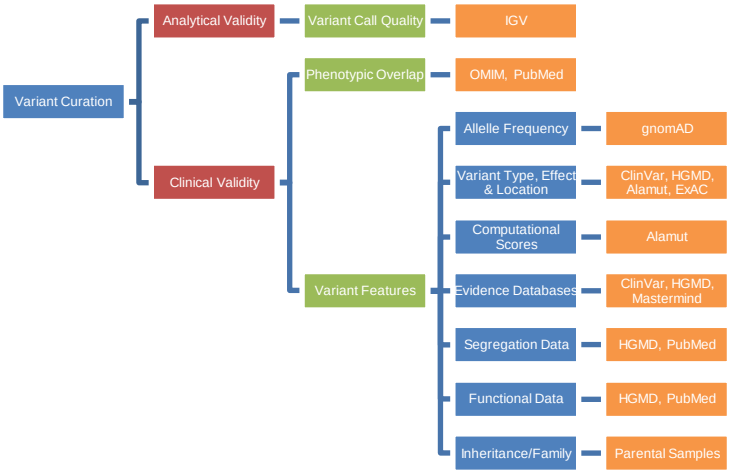
- **Variant Prioritization**
- **Variant Curation**
  - Evidence Types
  - Databases and Tools
- **Variant Classification**
  - ACMG criteria
- **Orthogonal Confirmation and Identity Testing**
- **Return of Results**

Variant Prioritization

- Process of ranking and prioritizing variants with respect to given phenotypes
- Variant prioritization tools (e.g., Phevor or Moon) combine lists of phenotypes (as HPO terms) with list of genetic variants (as vcf file) to generate a ranked list of annotated genetic variants potentially associated with disease
- Additional filters are applied to narrow down these ranked list of variants:
    - Quality (coverage, allelic balance)
    - Allele frequency (population databases)

Variant Curation

- Process of researching variants for evidence of pathogenicity



Variant Curation: Types of Evidence

- **Analytical validity**: quality of the variant call in the original sequencing reads
- **Phenotypic overlap**: degree of overlap between the established genetic condition and the patient's primary/presenting phenotypes
- **Allele frequency**: frequency of the genetic variant in the normal population
- **Variant type, effect, and location**: consequence of the genetic variant on the gene product
- **Computational scores**: in silico predictions based on evolutionary conservation or impact on structure/function of the gene product
- **Evidence databases**: previous classification of genetic variant in databases
- **Segregation or functional data**: previous knowledge of genetic variant in cohort or functional studies in the scientific and medical literature
- **Inheritance/Family**: inheritance pattern of the genetic variant (if known)

Variant Curation: Databases & Tools

Databases

- OMIM
- ClinVar
- HGMD
- Mastermind
- gnomAD & ExAC
- PubMed

Tools

- IGV
- Alamut Visual
- UCSC Genome Browser

Variant Curation: Databases & Tools

- **OMIM**: comprehensive online catalog of human genes and genetic disorders
- **ClinVar**: public database of classified genetic variants (with evidence)
- **HGMD**: private database of published genetic variants (with citations)
- **Mastermind**: private search engine for human disease genes and variants
- **gnomAD & ExAC**: public databases of genomes and exomes (including AF)
- **PubMed**: public resource for biomedical literature
- **IGV**: visualization tool for genomic data (including WGS and WES)
- **Alamut Visual**: variant analysis software (conservation, domains, predictions)
- **UCSC Genome Browser**: public platform to visualize “tracks” of genomic data

Variant Classification

- Process of scoring variants for pathogenicity according to a set of standard guidelines (e.g., ACMG)
- Five variant classification tiers:
  - **Pathogenic**
  - **Likely pathogenic**
  - **Variant of uncertain significance**
  - **Likely benign**
  - **Benign**

ACMG Criteria for Variant Interpretation

- The American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) published standards and guidelines for the interpretation of sequence variants in 2015
- The recommendations describe a process for classifying variants into five categories (P, LP, VUS, LB, B) based on criteria using typical types of variant evidence (e.g., population data, computational data, functional data, segregation data)
- For a given variant, the user selects the criteria based on the evidence observed for the variant
- Then, the criteria are combined according to the scoring rules to choose a classification from the five-tier system
- To provide flexibility, some criteria can be moved to another weight (higher or lower) using professional judgement depending on the evidence collected
- If a variant does not fulfill criteria for P/LP or B/LB or the evidence is conflicting, then the variant defaults to VUS

Criteria for Classifying P & LP Variants (1/3)

Evidence	Code	Evidence Type	Category
Very Strong	PVS1	Variant Effect	Null variant (nonsense, frameshift, canonical splice sites, initiation codon, single or multi-exon deletion) in a gene where LOF is a known mechanism of disease
Strong	PS1	Variant Location	Same amino change as a previously established pathogenic variant regardless of nucleotide change
Strong	PS2	Inheritance/Family	De novo (both maternity and paternity confirmed) in a patient with the disease and no family history
Strong	PS3	Functional Data	Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product
Strong	PS4	Evidence Databases	The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls

ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

Criteria for Classifying P & LP Variants (2/3)

Evidence	Code	Evidence Type	Category
Moderate	PM1	Variant Location	Located in a mutational hot spot and/or critical and well-established functional domain (i.e., active site of an enzyme) without benign variation
Moderate	PM2	Allele Frequency	Absent from controls (or at extremely low frequency if recessive) in ESP, 1KG, ExAC, or gnomAD
Moderate	PM3	Inheritance/Family	For recessive disorders, detected in trans with a pathogenic variant
Moderate	PM4	Variant Effect	Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants
Moderate	PM5	Variant Location	Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before

ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

Criteria for Classifying P & LP Variants (3/3)

Evidence	Code	Evidence Type	Category
Supporting	PP1	Segregation Data	Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease
Supporting	PP2	Variant Effect	Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease
Supporting	PP3	Computational Scores	Multiple lines of computational evidence support a deleterious effect on the gene or gene product
Supporting	PP4	Inheritance/Family	Patient's phenotype or family history is highly specific for a disease with a single genetic etiology
Supporting	PP5	Evidence Databases	Reputable source recently reports variant as pathogenic

ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

Criteria for Classifying B & LB Variants (1/2)

Evidence	Code	Evidence Type	Category
Stand-Alone	BA1	Allele Frequency	Allele frequency is > 5% in ESP, 1KG, ExAC, or gnomAD
Strong	BS1	Allele Frequency	Allele frequency is greater than expected for disorder
Strong	BS2	Allele Frequency	Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age
Strong	BS3	Functional Data	Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing
Strong	BS4	Segregation Data	Lack of segregation in affected members of a family

ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

Criteria for Classifying B & LB Variants (2/2)

Evidence	Code	Evidence Type	Category
Supporting	BP1	Variant Effect	Missense variant in a gene for which primarily truncating variants are known to cause disease
Supporting	BP2	Inheritance/Family	Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern
Supporting	BP3	Variant Effect	In-frame deletions/insertions in a repetitive region without a known function
Supporting	BP4	Computational Scores	Multiple lines of computational evidence suggest no impact on the gene or gene product
Supporting	BP5	Inheritance/Family	Variant found in a case with an alternate molecular basis for disease
Supporting	BP6	Evidence Databases	Reputable source recently reports variant as benign
Supporting	BP7	Variant Effect	A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved

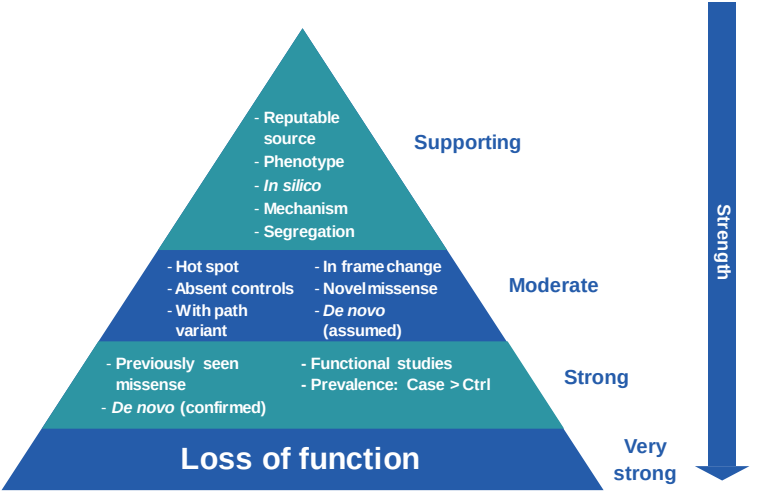
ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

Rules for Combining Criteria (P/LP or B/LB)

ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

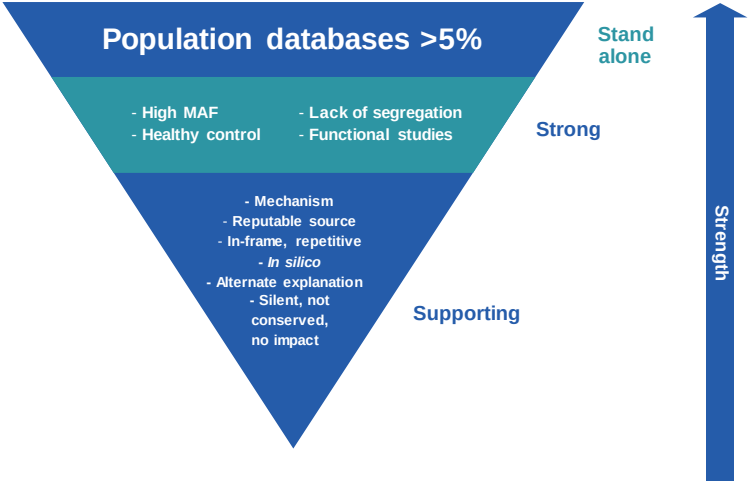
Pathogenic	(i) 1 Very strong (PVS1) *AND* (a) ≥1 Strong (PS1–PS4) *OR* (b) ≥2 Moderate (PM1–PM6) *OR* (c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) *OR* (d) ≥2 Supporting (PP1–PP5) (ii) ≥2 Strong (PS1–PS4) *OR* (iii) 1 Strong (PS1–PS4) *AND* (a)≥3 Moderate (PM1–PM6) *OR* (b)2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR* (c)1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Likely pathogenic	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR* (ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR* (iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR* (iv) ≥3 Moderate (PM1–PM6) *OR* (v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR* (vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR* (ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) *OR* (ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	(i) Other criteria shown above are not met OR (ii) the criteria for benign and pathogenic are contradictory

Evidence of Pathogenic Impact



ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

Evidence of Benign Impact



ACMG. Standards and guidelines for the interpretation of sequence variants. Genetics in Medicine. 2015.

ACMG Criteria: Special Considerations

- It is **not** appropriate to use the ACMG classification rules for variants in a gene of uncertain significance (GUS)
    - GUS's do not have established gene-disease associations in OMIM, but may exhibit strong phenotypic overlap with the patient as determined by published biomedical literature
    - Variants in GUS's should **always** be classified as VUS
- Caution should be exercised when using the ACMG classification rules to evaluate variants in healthy or asymptomatic individuals or to interpret secondary findings
- The ACMG classification rules should **not** be applied to variants associated with pharmacogenomics, common or complex disorders, or somatic variants

Orthogonal Confirmation & Identity Testing

- SNVs and indels can be confirmed by Sanger Sequencing
- CNVs (deletions and duplications) can be confirmed by multiplexed ligation-dependent probe amplification (MLPA)
- Samples can be tested for familial relationships by STR analysis
- STR identity check is useful for determining the inheritance or phasing of a particular variant(s)

Return of Results

- Genomic test reports typically include:
  - Patient, ordering, and sample information
  - Test results (variant, classification, genetic conditions, confirmation, analysis of parental samples)
  - Primary findings (variant and gene information)
    - Related to the primary or presenting phenotype(s)
  - Incidental findings (variant and gene information)
    - Not related to the primary or presenting phenotype(s)