Introduction to Population Genomics: Scope and Applications

@BeGenomics

Basic terms used in population Genomics

> **Single Nucleotide Polymorphism (SNP):** SNPs are variations in a DNA sequence that occur when a single nucleotide (A, T, C, or G) differs between individuals.

For example, a A nucleotide present at a specific location in a reference genome may be replaced by an G in a minority of individuals.

The two possible nucleotide variations of this SNP - A or G are called alleles.

> **Allele**

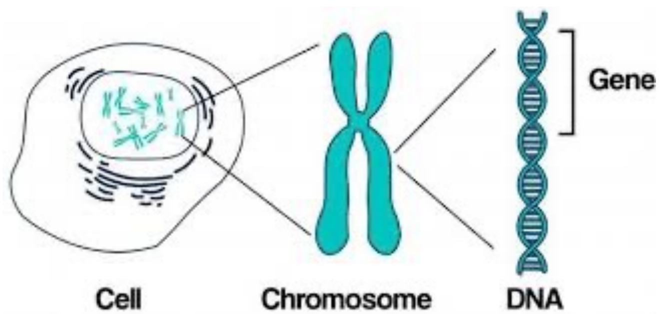
The **possible variants** of the SNP are called **alleles**.

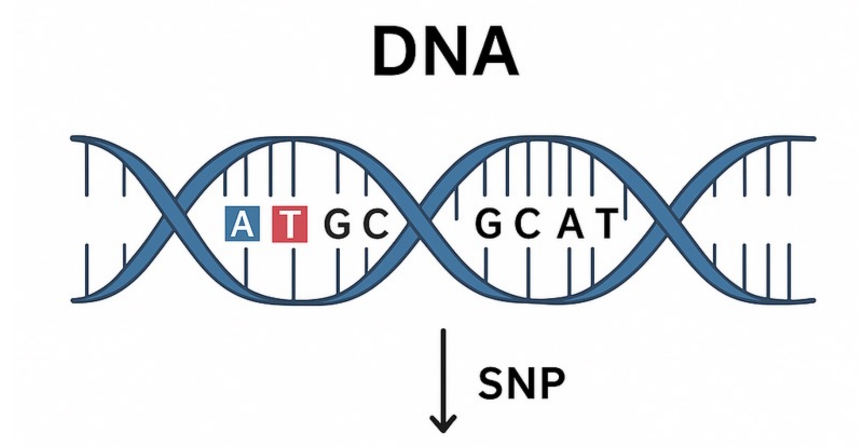
In diploid species like humans, two full sets of chromosomes are present, meaning each individual has two alleles for any given gene.

If both alleles are the same, the genotype is referred to as **homozygous**. If the alleles are different, the genotype is referred to as **heterozygous**.

> **Genotype**

- A genotype is the genetic makeup of an organism.
- It represents the specific combination of alleles an individual possesses for a particular gene or across their entire genome.
- Genotypes can be homozygous (having two identical alleles) or heterozygous (having two different alleles).
- Genotypes can also be represented by the actual DNA sequence at a specific location, such as AA, AG, GG, TT.
- For instance, a genotype for eye color could be ”GG" (homozygous dominant for brown eyes), ”gg" (homozygous recessive for blue eyes), or ”Gg" (heterozygous, where brown is dominant).
- **Alleles are the building blocks of a genotype**
- The genotype is determined by the specific combination of alleles an individual inherits.
- The genotype, in turn, influences the phenotype of an organism





AAGGCCCCATGC**A**GCATCTGGAAGT

Allele 1: AAGGCCCCATGCAGCATCTGGAAGT
Allele 2: AAGGCCCCATGCGGCATCTGGAAGT

A/A   G/G   A/G

Genotype

Possible genotypes at this SNP:

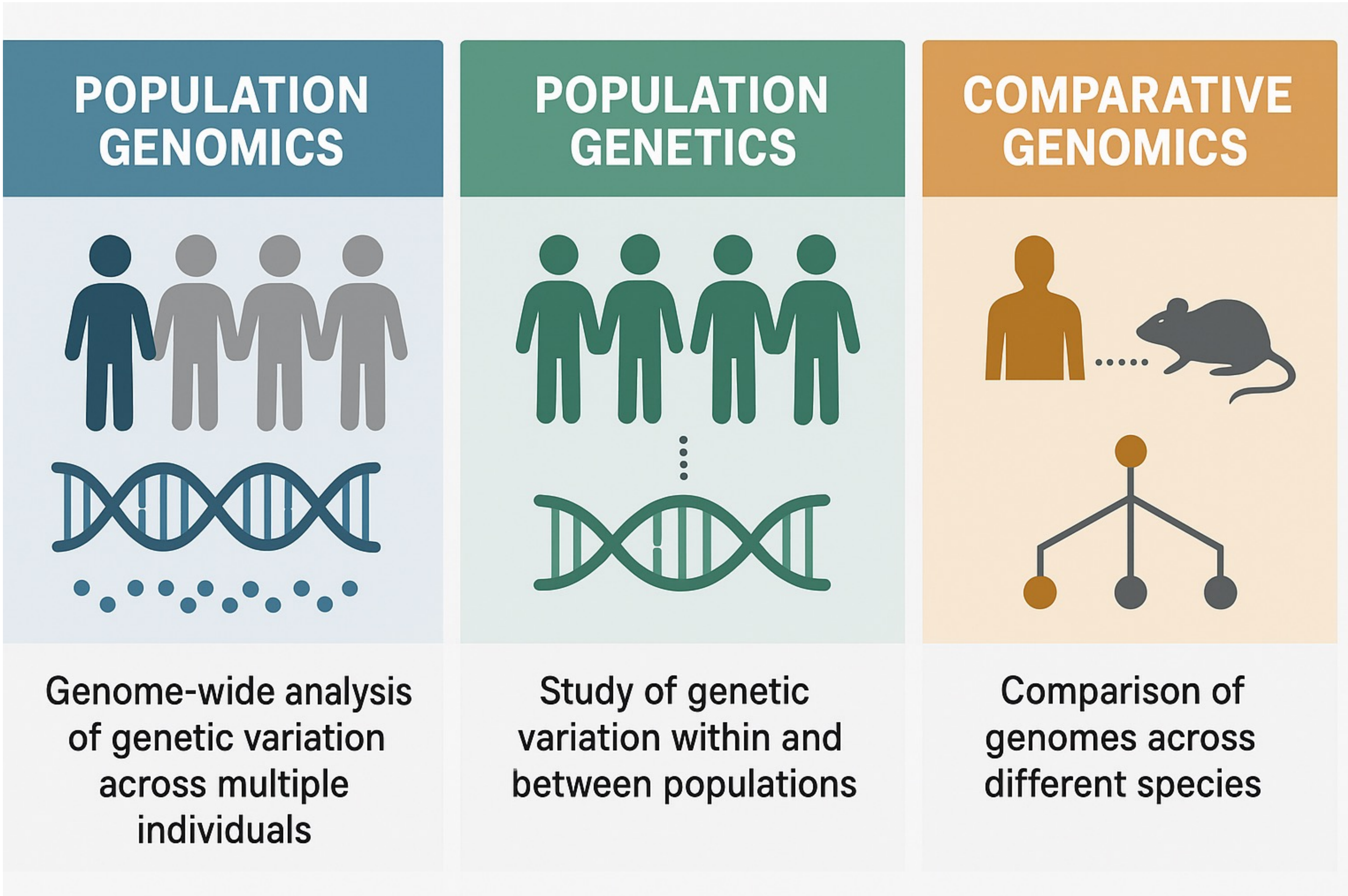
- **A/A** (homozygous for A)
- **G/G** (homozygous for G)
- **A/G** (heterozygous, carrying both alleles)

Introduction to Population Genomics

- Population genomics examines the genetic makeup of many individuals to understand **variation**, adaptation, and evolution.
- **Focus:** Uses genome-wide data instead of single genes, providing a holistic view of population genetics.
- **Key Goals:**
- Identify genetic diversity within and between populations.
- Understand evolutionary processes such as selection, drift, and migration.
- Link genetic variation to traits, diseases, and environmental adaptation.
- **Data Sources:** Whole-genome sequencing, SNP arrays, multi-omics datasets.

From Population Genetics to Population Genomics

- Population Genetics: Studies limited genetic markers
- Population Genomics: Genome-wide analysis for deeper insights
- **Few genetic markers** to **millions** across genome


POPULATION GENOMICS
Genome-wide analysis of genetic variation across multiple individuals
POPULATION GENETICS
Study of genetic variation within and between populations
COMPARATIVE GENOMICS
Comparison of genomes across different species

Genetic Variation

- Genetic variation = differences in DNA sequences among individuals in a population

Sources of variation:

- **Mutations** (e.g., SNPs, insertions, deletions)
- **Recombination** during meiosis
- **Gene flow** (migration introducing new alleles)

Essential for:

- **Adaptation and evolution**
- **Disease resistance and susceptibility**
- **Maintaining biodiversity**

Types of Genetic Variation

Types of variation:

- Single nucleotide polymorphisms (**SNPs**)
- **Structural variants** (CNVs, inversions, duplications)
- **Microsatellites** / short tandem repeats

Measurement in population genetics:

- **Allele frequency** (p, q)
- **Genotype frequency**
- **Heterozygosity** (expected vs observed)
- **Hardy–Weinberg equilibrium** as a baseline model

Scope of Population Genomics

Understanding Genetic Diversity

- Cataloguing genome-wide variation across populations
- Identifying rare and common variants

Evolutionary Insights

- Reconstructing demographic history & migrations
- Detecting natural selection and adaptation

Population Structure

- Differentiation between subpopulations
- Gene flow and admixture patterns

Scope of Population Genomics

Medical Genomics

- Identifying disease-associated variants
- Precision medicine tailored to population-specific risks

Agriculture & Conservation

- Improving crop and livestock breeding strategies
- Conserving endangered species via genetic monitoring

Human History & Anthropology

- Tracing ancestry and evolutionary relationships
- Linking genomics with archaeological findings

Tools and Methods in Population Genomics

- **DNA Sequencing Platforms:**
- Whole-Genome Sequencing (WGS) – captures all genetic variation.
- Whole-Exome Sequencing (WES) – focuses on coding regions.
- Targeted Sequencing Panels – specific genes of interest.
- **Genotyping Arrays:** High-throughput SNP detection for population-scale studies.
- **Bioinformatics Pipelines:** Tools like GATK, bcftools, PLINK for variant calling, filtering, and analysis.
- **Databases & Repositories:**
- 1000 Genomes Project, gnomAD, Ensembl, dbSNP for reference data.
- **High-Performance Computing (HPC):** Needed for processing large genomic datasets.

Methods in Population Genomics

- **Allele & Genotype Frequency Analysis:** Calculate and compare across populations.
- **Population Structure Analysis:** PCA, STRUCTURE, ADMIXTURE to identify genetic clusters.
- **Linkage Disequilibrium Mapping:** Detect non-random allele associations.
- **Demographic Inference:** Estimate migration, bottlenecks, and population size changes using coalescent models.
- **Genome-Wide Association Studies (GWAS):** Link genetic variants to phenotypic traits/diseases.

Challenges in Population Genomics

1. Technical Challenges

- **Data Volume & Complexity:** Whole-genome datasets can reach petabyte scale; require advanced storage and computing.
- **Sequencing Errors & Bias:** Platform-specific artifacts can affect variant calling accuracy.
- **Incomplete Reference Genomes:** Many populations are underrepresented in reference data.

2. Analytical Challenges

- **Population Structure & Admixture:** Confounding effects in association studies.
- **Statistical Power:** Large sample sizes needed for rare variant detection.
- **Interpretation of Variants:** Many variants of unknown significance (VUS) remain.

3. Ethical & Social Challenges

Application and Future Directions

- **Public Health:** Tracking pathogen evolution, drug resistance, and genetic epidemiology.
- **Agriculture & Breeding:** Enhancing crop and livestock traits using population genomics.
- **Conservation Biology:** Managing genetic diversity in endangered species.
- **AI & Big Data Integration:** Leveraging machine learning for large-scale genomic datasets.
- Integration of multi-omics (genomics, transcriptomics, epigenomics)
- **Precision Medicine:** Tailoring diagnostics and treatments to population-specific genetic profiles.

What is Personalized Medicine?

- Tailoring medical treatment to an individual's **genetic makeup, lifestyle, and environment**
- Moves away from "one-size-fits-all" → toward **patient-specific therapy**
- Key enablers:
- **Genomic sequencing** (e.g., whole-genome, exome, targeted panels)
- **Biomarkers** (DNA, RNA, proteins, metabolites)
- **Big data & AI** for integration and prediction
- Goal: **right drug, right dose, right patient, right time**

Applications of Personalized Medicine

- **Oncology**: targeted therapies based on tumor mutations (e.g., EGFR, BRCA)
- **Pharmacogenomics**: optimizing drug response & reducing adverse effects
- **Rare disease diagnosis**: genome sequencing for unexplained syndromes
- **Cardiovascular & metabolic diseases**: identifying genetic risk & preventive strategies
- **Public health**: stratifying populations for precision prevention
- **Benefits:** improved outcomes, reduced side effects, cost-effective healthcare

Summary

- Population genomics provides a comprehensive view of genetic variation
- Supports diverse applications from evolution to medicine
- Requires advanced tools and ethical considerations
- Future lies in multi-omics integration and AI-driven insights

Thanks

- Questions?

- https://www.begenomics.com/genomichub