

# Molecular Breeding Data Analytics

## Course Content

### Module I : Molecular Breeding Data Handling and Analysis

#### Basics of Molecular Markers, QTL mapping, Marker-Assisted Selection (MAS)

- Overview of different types of molecular markers: SSR, SNP, AFLP, etc.
- Principles and workflow of marker-assisted selection in crop improvement.
- Applications of molecular markers in trait mapping and breeding programs.
- QTL Mapping-Linkage and Association Mapping, Linkage disequilibrium.

#### Genotyping of SSR Markers, Scoring and Data Generation

- Experimental design and protocol overview for SSR genotyping.
- Methods for allele scoring from gel or capillary electrophoresis.
- Preparing data matrices for downstream analysis.
- Genetic Diversity Studies; Estimating allelic diversity, heterozygosity, and polymorphism information content (PIC), Construction of dendrograms

#### Population Structure Analysis using STRUCTURE Software

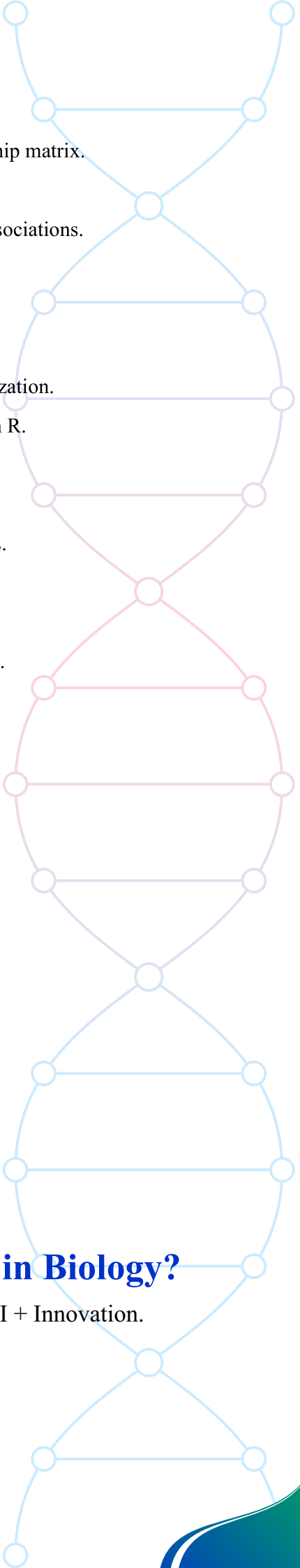
- Introduction to Bayesian clustering methods for determining genetic structure.
- Preparation of input files, setting parameters, and interpreting STRUCTURE outputs.
- Visualization of population clusters and admixture using graphical tools.

#### SNP Data Analysis: HapMap Format, Imputation, and Filtering

- Understanding SNP data formats including HapMap and VCF.
- Quality control steps: filtering by MAF, missing data, and heterozygosity.
- Imputation of missing genotypes and preparing datasets for GWAS.

### Module II : GWAS, Haplotypes Study & KASP Assay

#### GWAS using TASSEL Software

- 
- Introduction to GWAS principles: population structure correction, kinship matrix.
  - Step-by-step execution of GWAS in TASSEL using example datasets.
  - Interpreting Manhattan plots, Q-Q plots, and significant marker-trait associations.

### **GWAS using GAPIT Package in R Studio**

- Installing and working with the GAPIT R package for GWAS.
- Data formatting, model selection (GLM, MLM, FarmCPU), and visualization.
- Best practices for reproducibility and automation of GWAS pipelines in R.

### **Candidate Gene-Based Association Studies**

- Targeted association studies using known genes related to specific traits.

### **Candidate Gene Identification**

- Strategies to pinpoint potential candidate genes near associated markers.
- Utilizing genome browsers, gene annotations, and functional databases.
- Prioritizing candidate genes for validation and functional studies.

### **Haplotype Analysis**

- Definition and significance of haplotypes in genetic analysis.
- Identification of haplotype with in linkage disequilibrium (LD) region.
- Association of haplotype with phenotypic traits

### **KASP Assay for SNP Marker Genotyping**

- Overview of Kompetitive Allele Specific PCR (KASP) technology.
- Designing KASP assays for SNP markers and trait-linked alleles.
- KASP data analysis and data interpretations.

## **Ready to Build Real-World AI Solutions in Biology?**

Join the next generation of researchers combining Biology + AI + Innovation.

**APPLY NOW**

Visit Website : [www.omniedgesci.com](http://www.omniedgesci.com)