Advanced Statistical Genetics Training Course

Professional Development Training Course Outline

Category	Biotechnology and Pharmaceutical Development	**Duration**	10 Days
Base Fee	$4,000 USD	**Delivery Mode**	Online / On-site Hybrid

Course Introduction

Statistical Genetics and Genomics are rapidly evolving fields at the intersection of Big Data, biology, and advanced computation, driving breakthroughs in Precision Medicine and agricultural innovation. The ability to effectively analyze massive, high-dimensional Multi-Omics datasets is critical for dissecting the complex genetic architecture of human diseases and economically important traits in plants and livestock. Advanced Statistical Genetics Training Course provides participants with deep expertise in state-of-the-art Quantitative Genetics and Bioinformatics methodologies, focusing on modern statistical modeling, Causal Inference, and the application of Machine Learning techniques. Trainees will master complex analytical pipelines to transition from raw data to actionable biological insights, addressing challenges like population structure, rare variants, and the integration of diverse molecular data for comprehensive Trait Prediction.

This intensive program emphasizes hands-on practical skills using leading computational tools like R/Bioconductor and PLINK, ensuring graduates are immediately proficient in implementing advanced methods such as Genome-Wide Association Studies (GWAS), Polygenic Risk Scores (PRS), Mendelian Randomization (MR), and Single-Cell Omics analysis. By focusing on cutting-edge techniques and real-world Case Studies, the course prepares researchers and data scientists to lead impactful studies in diverse sectors, from pharmaceutical research for Drug Target Identification to advanced breeding programs for Climate Resilience. Completion of this training signifies mastery of the advanced quantitative methods required to navigate the future landscape of Genetic Data Science and contribute to significant scientific and organizational advancements.

Programme Curriculum

Advanced Statistical Genetics Training Course

Introduction

Statistical Genetics and Genomics are rapidly evolving fields at the intersection of Big Data, biology, and advanced computation, driving breakthroughs in Precision Medicine and agricultural innovation. The ability to effectively analyze massive, high-dimensional Multi-Omics datasets is critical for dissecting the complex genetic architecture of human diseases and economically important traits in plants and livestock. Advanced Statistical Genetics Training Course provides participants with deep expertise in state-of-the-art Quantitative Genetics and Bioinformatics methodologies, focusing on modern statistical modeling, Causal Inference, and the application of Machine Learning techniques. Trainees will master complex analytical pipelines to transition from raw data to actionable biological insights, addressing challenges like population structure, rare variants, and the integration of diverse molecular data for comprehensive Trait Prediction.

This intensive program emphasizes hands-on practical skills using leading computational tools like R/Bioconductor and PLINK, ensuring graduates are immediately proficient in implementing advanced methods such as Genome-Wide Association Studies (GWAS), Polygenic Risk Scores (PRS), Mendelian Randomization (MR), and Single-Cell Omics analysis. By focusing on cutting-edge techniques and real-world Case Studies, the course prepares researchers and data scientists to lead impactful studies in diverse sectors, from pharmaceutical research for Drug Target Identification to advanced breeding programs for Climate Resilience. Completion of this training signifies mastery of the advanced quantitative methods required to navigate the future landscape of Genetic Data Science and contribute to significant scientific and organizational advancements.

Course Duration

10 days

Course Objectives

Upon completion, participants will be able to:

1. Master advanced statistical modeling for complex trait analysis.
2. Design, execute, and critically appraise Genome-Wide Association Studies (GWAS), including robust Quality Control (QC).
3. Implement methods for detecting and correcting for Population Structure and cryptic relatedness.
4. Develop and validate highly accurate Polygenic Risk Scores (PRS) and Genomic Prediction models.
5. Apply Mendelian Randomization (MR) for Causal Inference in observational genetic data.
6. Analyze rare variants and perform gene-set/pathway enrichment analysis.
7. Integrate diverse Multi-Omics data using advanced statistical frameworks.
8. Perform eQTL (expression Quantitative Trait Loci) and other regulatory element mapping.
9. Utilize Machine Learning and Deep Learning techniques for high-dimensional genetic data prediction.
10. Analyze data from Next-Generation Sequencing (NGS) platforms, including Single-Cell Genomics.
11. Develop, benchmark, and apply custom Bioinformatics Pipelines for genetic data processing.
12. Interpret and visualize results effectively for Drug Target Identification and clinical translation.
13. Practice Reproducible Research through version control and standardized reporting.

Target Audience

1. Bioinformaticians and Computational Biologists.
2. Geneticists and Epidemiologists.
3. Data Scientists and Statisticians.
4. R&D Scientists in Pharmaceutical and Biotechnology companies.
5. Researchers in Plant and Animal Breeding.
6. Postdoctoral Researchers and PhD Students.

7. Medical Researchers.
8. Biomedical Analysts.

Course Modules

Module 1: Foundational Statistical and Population Genetics Review

- Advanced concepts of Linkage Disequilibrium (LD) and Haplotype phasing.
- Review of complex inheritance models and quantitative genetics theory
- In-depth treatment of Hardy-Weinberg Equilibrium deviations and population subdivision
- Mathematical basis of linear mixed models (LMMs) for relatedness correction.
- Bayesian versus frequentist approaches in genetic parameter estimation.
- Case Study: Estimating the heritability of height in different human populations using GCTA.

Module 2: Advanced Genome-Wide Association Studies (GWAS)

- Rigorous sample and variant Quality Control (QC) using PLINK and Hail.
- Handling covariates, cryptic relatedness, and population structure
- Meta-analysis techniques and heterogeneity assessment.
- Statistical power calculation and effect size estimation
- Conditional and joint association analysis
- Case Study: Meta-analysis of multiple cohorts to identify novel loci for Type 2 Diabetes.

Module 3: Imputation and Fine-Mapping

- Principles of genotype Imputation and use of reference panels
- Software training for imputation
- Statistical methods for Fine-Mapping causal variants
- Credible set identification and visualization of fine-mapping results.
- Accounting for linkage disequilibrium structure in causal variant prioritization.
- Case Study: Fine-mapping a coronary artery disease locus to pinpoint the likely causal SNP.

Module 4: Analysis of Rare and Structural Variants

- Burden and Variance Component Tests for Rare Variants
- Analyzing data from whole-exome and whole-genome sequencing
- Statistical challenges and methods for analyzing Structural Variants
- Gene-based and pathway-based testing methodologies.
- Collapsing strategies for aggregating rare variants within functional units.
- Case Study: Identifying a novel disease gene for intellectual disability using WES data and SKAT-O.

Module 5: Polygenic Risk Scores (PRS) and Genomic Prediction

- Theoretical foundations of Genomic Selection and Genomic Best Linear Unbiased Prediction
- Methods for developing Polygenic Risk Scores
- Strategies for cross-population PRS prediction and transportability challenges.
- Model evaluation metrics and calibration for clinical utility
- Practical implementation of PRS in R and high-performance computing environments.
- Case Study: Building and validating a PRS for breast cancer risk in a diverse biobank cohort.

Module 6: Mendelian Randomization (MR) and Causal Inference

- Core assumptions and limitations of the Mendelian Randomization framework.
- Two-sample MR methods and sensitivity analyses.
- Advanced MR techniques for pleiotropy and heterogeneity
- Mediation analysis and multi-variable MR for distinguishing independent causal effects.
- Applying MR for validating Drug Targets and understanding disease mechanisms.
- Case Study: Using MR to test the causal effect of LDL cholesterol on Alzheimer's disease risk.

Module 7: Transcriptome-Wide Association Studies (TWAS)

- Principles of integrating genetic data with Gene Expression data.
- Statistical models for predicting gene expression from genotypes
- Implementation of TWAS methodologies to identify risk genes.
- Distinguishing between TWAS and GWAS findings for biological interpretation.
- Application of TWAS to understand tissue-specific effects.
- Case Study: Identifying expression-driven associations in schizophrenia using brain eQTL data.

Module 8: Introduction to Statistical Multi-Omics Integration

- Conceptual framework for integrating diverse data types
- Data normalization, batch correction, and harmonization across Omics layers.
- Statistical methods for data fusion
- Network and pathway analysis using integrated genetic findings.
- Interpreting and visualizing multi-omic association signals.
- Case Study: Integrating GWAS and Metabolomics data to uncover novel pathways related to cardiovascular health.

Module 9: Statistical Methods for Epigenetics and Chromatin Data

- Analysis of DNA Methylation data and quality control.
- Statistical modeling of ATAC-seq and ChIP-seq for regulatory element analysis.
- Genetics of methylation and chromatin accessibility
- Methods for linking regulatory variants to target genes
- Differential analysis of epigenetic marks and environmental effects.
- Case Study: Linking GWAS hits to target genes by integrating caQTL data in immune cells.

Module 10: Introduction to Single-Cell Genomics Analysis

- Single-Cell RNA-seq data processing and normalization.
- Dimensionality reduction and clustering for cell type identification.
- Differential gene expression and trajectory inference in single cells.
- Statistical challenges of sparse and highly variable single-cell data.
- Introduction to single-cell eQTL and rare cell type analysis.
- Case Study: Identifying novel sub-populations of T-cells associated with autoimmune disease using scRNA-seq.

Module 11: Machine Learning and AI in Statistical Genetics

- Overview of supervised and unsupervised Machine Learning techniques
- Application of ML for complex trait Trait Prediction and classification.
- Introduction to Deep Learning architectures for sequence and imaging genetics.
- Feature selection and interpretability in ML genetic models

- Cross-validation and robust model deployment strategies.
- Case Study: Utilizing a Convolutional Neural Network to predict regulatory element function from DNA sequence.

Module 12: Ethical, Legal, and Social Implications (ELSI) and Data Sharing

- Ethical frameworks for large-scale human genetic data collection and use.
- Considerations for data privacy, de-identification, and security
- Informed consent for genomic research and sharing
- Statistical biases in genetic studies and addressing issues of health equity.
- Best practices for Reproducible Research and public data deposition
- Case Study: Debating the ethical implications of using PRS in a clinical setting for risk stratification.

Module 13: Statistical Methods for Functional Genomics

- Functional annotation of non-coding GWAS variants
- Integration of GWAS with regulatory genomics and protein-protein interaction networks.
- Bayesian methods for prioritization of causal genes and variants.
- Inference of gene regulatory networks from Omics data.
- Techniques for interpreting results in the context of pathways and biological processes.
- Case Study: Prioritizing causal genes for schizophrenia using Genomic Convergence methods.

Module 14: Statistical Genetics in Applied Contexts

- Genomic Selection and breeding value estimation in livestock and crops.
- Analysis of Quantitative Trait Loci in experimental crosses.
- Methods for detecting recent Natural Selection and local adaptation in populations.
- Statistical models for genotype-by-environment interaction.
- Application of statistical methods for de-extinction and conservation genetics.
- Case Study: Implementing a Genomic Selection program in dairy cattle to improve milk yield.

Module 15: Advanced Computational Tools and Pipelines

- Advanced R/Bioconductor and Python libraries for genomic analysis.
- Effective use of PLINK and GCTA scripting for large datasets.
- Cloud computing and High-Performance Computing best practices.
- Building Snakemake or Nextflow Bioinformatics Pipelines for automation.
- Containerization with Docker/Singularity for Reproducible Research.
- Case Study: Developing a fully automated GWAS to PRS pipeline using Nextflow on a cloud platform.

Training Methodology

The course employs an Active Learning approach, blending theoretical depth with intensive, Hands-on Practical Training.

- Lectures & Discussions.
- Practical Lab Sessions
- Case Study Analysis.
- Pipeline Development Workshop.
- Peer Review & Presentation.

Register as a group from 3 participants for a Discount

Send us an email: info@fineskilltrainingcenter.org or call +254769199797

Certification

Upon successful completion of this training, participants will be issued with a globally- recognized certificate.

Tailor-Made Course

We also offer tailor-made courses based on your needs.

Key Notes

a. The participant must be conversant with English.

b. Upon completion of training the participant will be issued with an Authorized Training Certificate

c. Course duration is flexible and the contents can be modified to fit any number of days.

d. The course fee includes facilitation training materials, 2 coffee breaks, buffet lunch and A Certificate upon successful completion of Training.

e. One-year post-training support Consultation and Coaching provided after the course.

f. Payment should be done at least a week before commence of the training, to Fineskill Training Center account, as indicated in the invoice so as to enable us prepare better for you.

Upcoming Scheduled Sessions

Start Date	End Date	Location	Price
23 Nov 2026	04 Dec 2026	Online	$2,000
23 Nov 2026	04 Dec 2026	Physical	$3,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$6,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$9,000
23 Nov 2026	04 Dec 2026	Physical	$12,000
23 Nov 2026	04 Dec 2026	Physical	$12,000

Ready to enroll or request a customized program? Book online or contact us:

Register Online Now

Email: info@fineskilltrainingcenter.com | Website: www.fineskilltrainingcenter.com