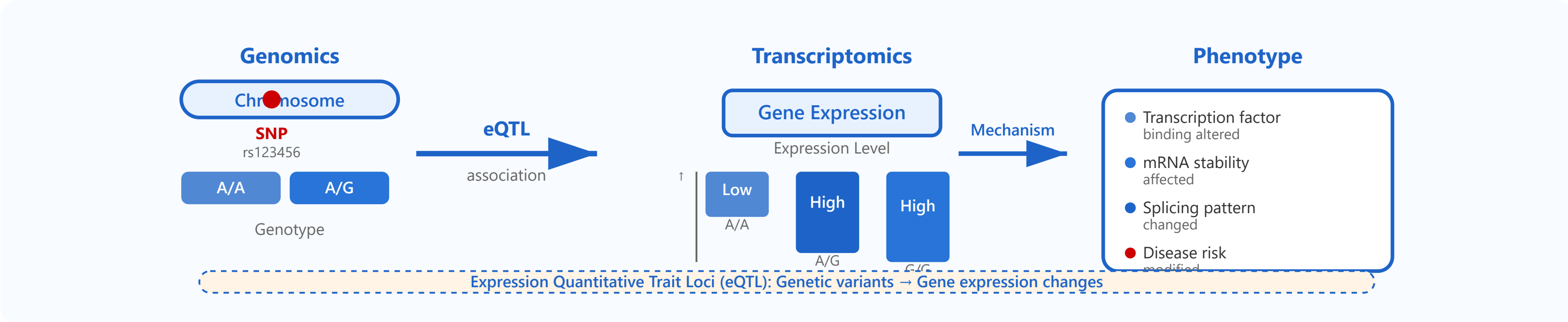
Genomics + Transcriptomics Integration



eQTL Analysis

Expression quantitative trait loci mapping

ASE Detection

Allele-specific expression patterns

Splicing QTLs

Genetic variants affecting RNA splicing

Regulatory Variants

Non-coding variants and gene expression

Allele-specific Binding

Transcription factor binding affected by SNPs

Detailed Analysis of Integration Methods

1. eQTL Analysis (Expression Quantitative Trait Loci)

Definition: eQTLs are genomic loci that contribute to variation in gene expression levels. They represent genetic variants (typically SNPs) that are associated with changes in mRNA expression.

Types of eQTLs:

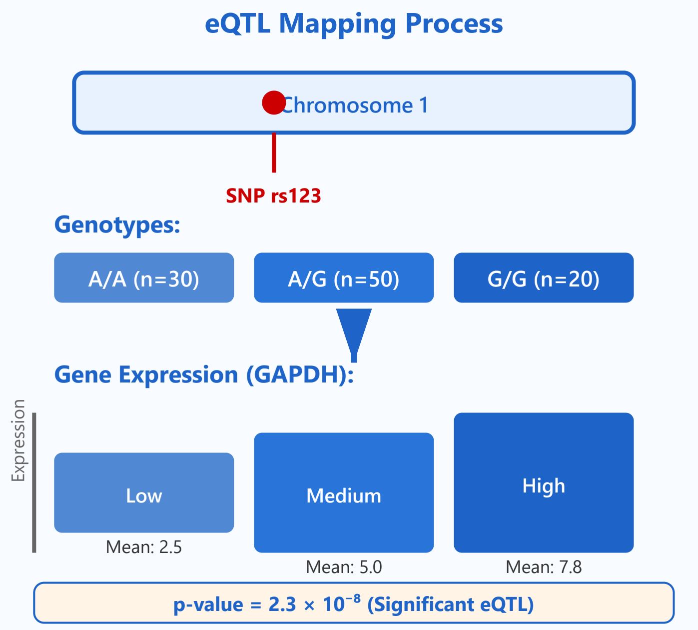
- **cis-eQTLs:** Variants located near the gene they regulate (typically within 1 Mb), acting locally on nearby genes
- **trans-eQTLs:** Variants located far from or on different chromosomes than the genes they regulate, acting through diffusible factors

Applications:

- Identifying regulatory mechanisms underlying GWAS signals
- Prioritizing causal genes at disease-associated loci
- Understanding tissue-specific regulation
- Drug target identification and validation

Statistical Approach:

Linear regression testing association between genotype and expression: Expression ~ Genotype + Covariates



Key Points:

- eQTL studies require both genotype and gene expression data from the same individuals
- Most eQTLs are cis-acting and tissue-specific
- eQTL data helps interpret GWAS findings by linking variants to gene regulation
- Multiple testing correction is essential due to testing millions of SNP-gene pairs

2. ASE Detection (Allele-Specific Expression)

Definition: ASE occurs when the two alleles of a gene are expressed at different levels in heterozygous individuals. This reveals cis-regulatory effects at the individual level.

Mechanisms Causing ASE:

- **Regulatory variants:** SNPs in promoters, enhancers, or UTRs affecting transcription
- **Genomic imprinting:** Parent-of-origin-specific expression
- **X-chromosome inactivation:** Random silencing of one X chromosome in females
- **Somatic mutations:** Acquired changes affecting expression

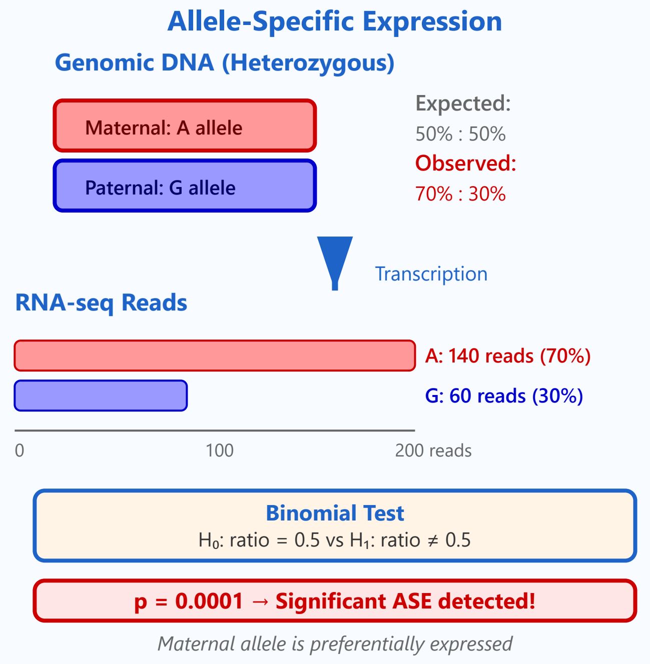
Detection Methods:

- RNA-seq with allelic read counting at heterozygous SNPs

- Statistical testing (binomial or beta-binomial models)
- Requires phasing to determine which allele is expressed more

Clinical Significance:

ASE can explain incomplete penetrance, variable expressivity, and personalized drug responses in genetic diseases.



Key Points:

- ASE provides direct evidence of cis-regulatory variation
- Requires sufficient RNA-seq read depth at heterozygous sites (typically >20 reads)
- Can identify imprinted genes and parent-of-origin effects
- Important for understanding disease mechanisms and personalized medicine

3. Splicing QTLs (sQTLs)

Definition: sQTLs are genetic variants that affect RNA splicing patterns, leading to different isoforms or exon usage patterns. They represent a major mechanism of gene regulation.

Types of Splicing Changes:

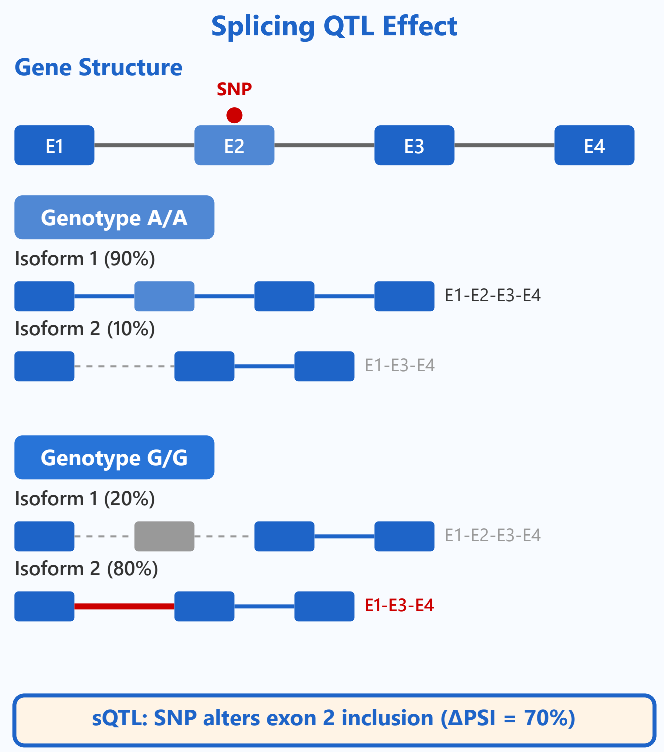
- **Exon skipping/inclusion:** Complete exons included or excluded
- **Alternative 5' or 3' splice sites:** Different splice site usage
- **Intron retention:** Failure to remove introns
- **Alternative first/last exons:** Different transcription start/end sites

Molecular Mechanisms:

- Disruption of splice site consensus sequences
- Creation or destruction of splicing enhancers/silencers
- Alteration of RNA secondary structure
- Changes in splicing factor binding sites

Detection Methods:

LeafCutter, MAJIQ, or rMATS for differential splicing analysis; testing association between genotype and junction usage or PSI (Percent Spliced In) values.



Key Points:

- sQTLs are highly prevalent and affect ~30% of genes
- Often independent of expression-level effects (eQTLs)
- Can create functionally distinct protein isoforms

- Important for understanding disease mechanisms and therapeutic targets
- Tissue-specific splicing patterns reflect cell-type-specific regulatory programs

4. Regulatory Variants in Non-Coding Regions

Definition: Regulatory variants are genetic variations in non-coding regions that affect gene expression by altering transcriptional regulation. Most disease-associated variants identified by GWAS are in non-coding regions.

Types of Regulatory Elements:

- **Promoters:** Regions directly upstream of genes controlling transcription initiation
- **Enhancers:** Distal elements that increase transcription (can be >1 Mb away)
- **Silencers:** Elements that repress transcription
- **Insulators:** Boundary elements that prevent enhancer-promoter interactions
- **3' UTRs:** Regions affecting mRNA stability and translation

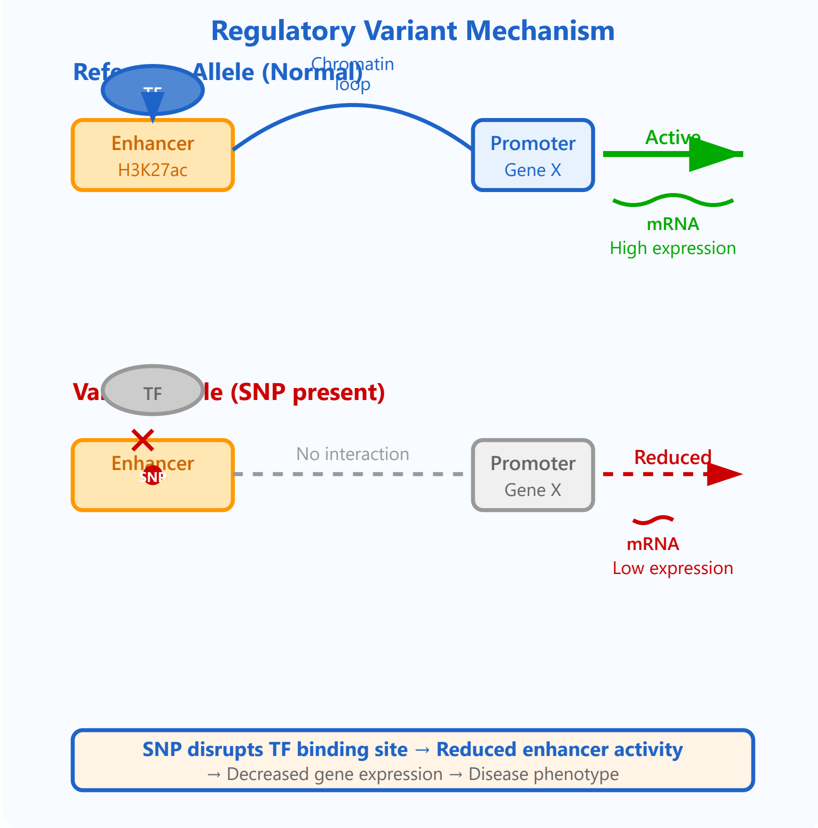
Identification Strategies:

- Integration with ChIP-seq data (histone marks, TF binding)

- ATAC-seq or DNase-seq for chromatin accessibility
- Chromosome conformation capture (Hi-C, ChIA-PET) for enhancer-promoter interactions
- Massively parallel reporter assays (MPRA)

Functional Consequences:

Altered transcription factor binding, changes in chromatin accessibility, disrupted enhancer-promoter loops, modified histone modifications.



Key Points:

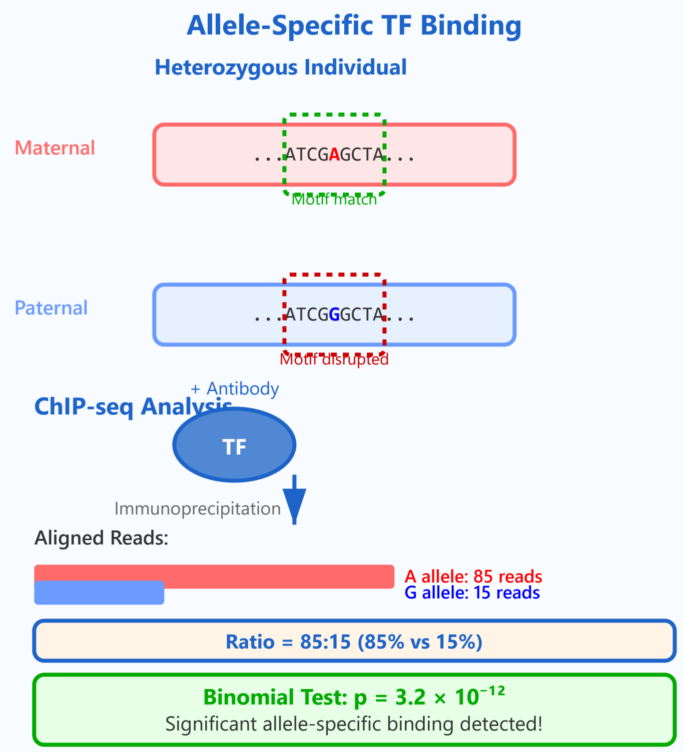
- ~90% of disease-associated variants from GWAS are in non-coding regions
- Enhancers can regulate genes over long distances (>1 Mb)
- Regulatory effects are often cell-type and tissue-specific
- Integration with epigenomic data is essential for interpretation
- Functional validation requires reporter assays or genome editing

5. Allele-Specific Transcription Factor Binding

Definition: Allele-specific binding (ASB) occurs when a transcription factor (TF) preferentially binds to one allele over another at heterozygous SNP sites. This is a direct mechanism linking genetic variation to gene regulation.

Detection Methods:

- **ChIP-seq:** Chromatin immunoprecipitation followed by sequencing to map TF binding
- **Analysis:** Count reads mapping to each allele at heterozygous SNPs within ChIP-seq peaks
- **Statistical test:** Binomial test or beta-binomial model for differential binding



Mechanisms:

- Direct disruption of TF binding motif by SNP
- Indirect effects through DNA shape changes
- Altered cooperativity with other TFs
- Changes in local chromatin structure

Biological Significance:

ASB events link genetic variants to gene expression changes and explain mechanisms of disease-associated variants. They provide functional evidence for causal variants in GWAS loci.

Key Points:

- ASB directly demonstrates functional impact of genetic variants on TF binding
- Requires sufficient ChIP-seq read depth at heterozygous sites (>20 reads)

- Can identify causal variants at GWAS loci by showing mechanism
- Often correlates with allele-specific expression (ASE) at target genes
- Useful for fine-mapping disease-associated variants and drug target discovery