Differential Expression

Statistical Models

Account for biological and technical variability

Negative Binomial

Models count data with overdispersion

Fold Change Thresholds

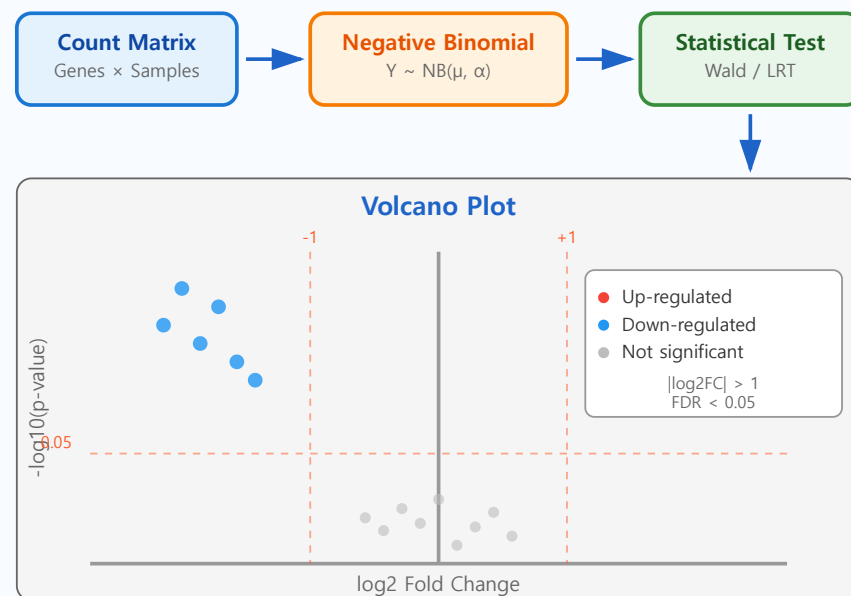Typically |log2FC| > 1 for biological significance

FDR Control

False Discovery Rate < 0.05 for multiple testing

Volcano Plots

Visualize FC vs statistical significance

DE Analysis Pipeline

Count Matrix — Genes × Samples → **Negative Binomial** — $Y \sim NB(\mu, \alpha)$ → **Statistical Test** — Wald / LRT

Volcano Plot

-1   +1

-log10(p-value)

0.05

- ● Up-regulated
- ● Down-regulated
- ● Not significant

|log2FC| > 1
FDR < 0.05

log2 Fold Change

💡 Balance statistical significance with biological relevance