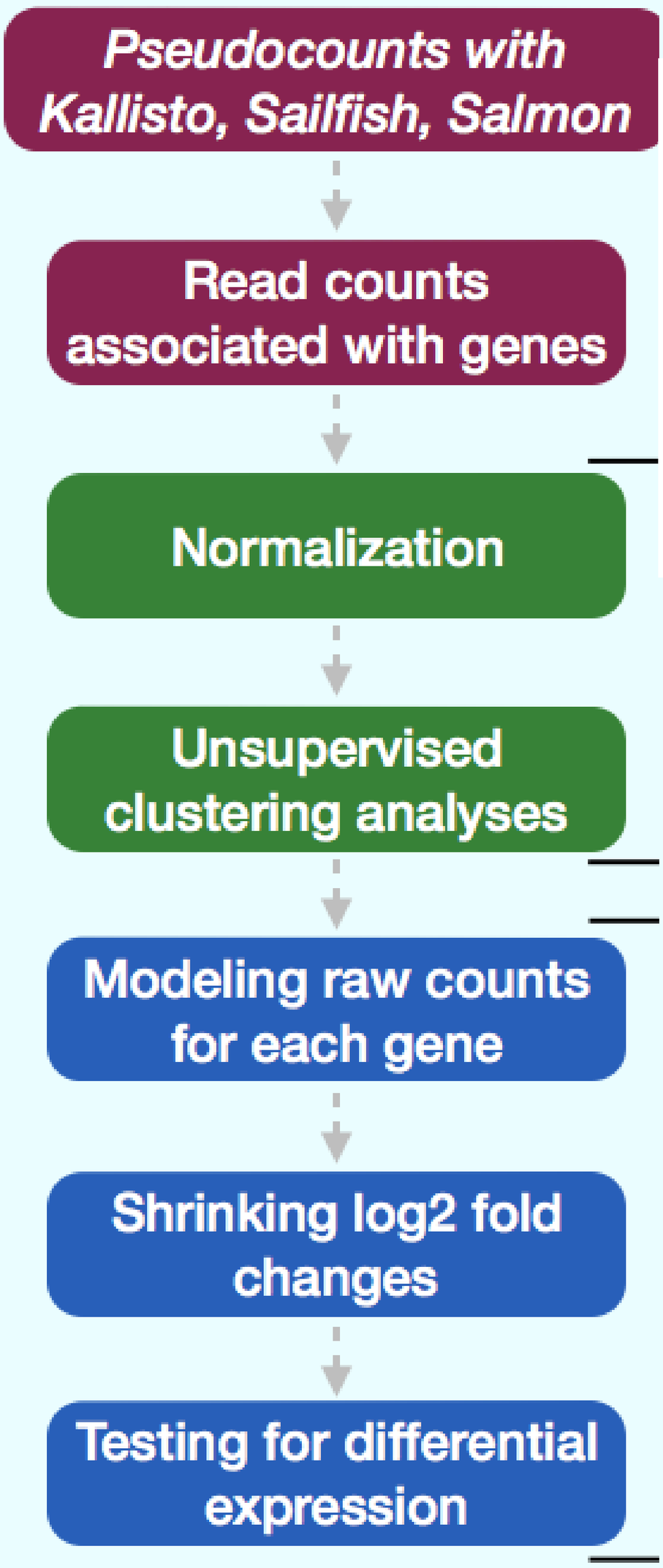
Python Port of DESEQ2

What is DESeq2?

DESeq2 is a tool used after alignment/quantification to analyze RNA-seq count data. Its goal is to find which genes change expression between conditions (e.g., treated vs. control). It normalizes sequencing depth, models counts with a negative binomial distribution, and identifies genes that are significantly up- or down-regulated.



Metric	Value
Size Factor Correlation	1.000000
Size Factor MAE	1.71e-15
Dispersion Correlation (log)	0.4949
Median Dispersion Ratio (Py/R)	0.68x
LFC Correlation (High Expr)	0.9530
Significant Gene Match (%)	98.7%
Top-500 DE Gene Overlap	80.4%
P-value Correlation	0.6897
LFC Sign Agreement	99.7%
LFC MAE	0.1102
LFC RMSE	0.3027
BaseMean Correlation	1.0000
Spearman (p-values)	0.9874
Spearman (|LFC|)	0.9595

DESeq2 in the RNA-Seq Pipeline

This diagram shows where DESeq2 fits within a typical RNA-seq analysis workflow. Upstream tools like Salmon, Kallisto, or STAR generate the count matrix. After loading these counts, DESeq2 performs normalization and statistical modeling. Clustering and PCA are often used for quality control, and DESeq2 then tests each gene for significant changes and applies log2 fold-change shrinkage.