

Perform differential expression analysis using gene expression values in...

Session ID: 58e556be • Date: February 03, 2026 at 03:57 PM • 2 messages • 1 analysis step(s)

YOU

Perform differential expression analysis using gene expression values in human_airway_smc_dexamethasone_treated_vs_ctrl.csv and its metadata file human_airway_smc_dexamethasone_treated_vs_ctrl_metadata.csv. Identify significantly differentially expressed genes using DESeq2. Generate relevant visualizations (e.g., MA plot, volcano plot, heatmap of top genes).

AGENT

☐ ****Analysis Complete****

☐ Dexamethasone treatment reveals 4,000 differentially expressed genes in airway muscle cells

****Key Findings:****

- 4,000 genes significantly differential (padj < 0.05)
- 2,193 genes upregulated, 1,807 genes downregulated
- Top gene: 4.57 log2 fold change, padj 7.41×10^{-132}
- 151 genes showed large upregulation effects
- 71% variance explained by first two PCA components

****Key Files:****

- ☐ `de_results.csv` (2.6 MB)
- ☐ `volcano_plot.png` (167.0 KB)
- ☐ `de_significant.csv` (504.2 KB)

****Output:**** 4 results, 5 plots, 1 reports

☐ Completed in 52.9s

☐ ***View the full report above for detailed methods and results.***

Results

Overview

Differential expression analysis was performed on human airway smooth muscle cell RNA-seq data comparing dexamethasone treatment versus control conditions using DESeq2. The dataset comprised 8 samples representing 4 control and 4 dexamethasone-treated samples across 4 different cell lines.

Data Quality and Processing

The initial dataset contained 63,677 genes across 8 samples. After applying low-count filtering (minimum 10 total counts), 22,369 genes remained for differential expression testing. Principal component analysis revealed high-quality data with clear biological signal, where PC1 captured 48% of variance and PC2 captured 23% of variance, totaling 71% of variance explained by the first two components.

Sample clustering showed clear separation between treatment groups along the PC1 axis, with control samples clustering on the left side and dexamethasone-treated samples clustering on the right side. The strong separation by treatment condition indicated robust differential gene expression response to dexamethasone treatment, with minimal technical variation as biological replicates clustered appropriately.

Main Findings

The differential expression analysis identified 4,000 significantly differentially expressed genes at an adjusted p-value threshold of 0.05. Of these significant genes, 2,193 were upregulated and 1,807 were downregulated in response to dexamethasone treatment. At more stringent significance thresholds, 2,901 genes remained significant at $p\text{-adj} < 0.01$ and 2,073 genes at $p\text{-adj} < 0.001$.

Analysis of effect sizes revealed 151 genes with large upregulation effects and 117 genes with large downregulation effects. The most significantly differentially expressed gene showed a log2 fold change of 4.57 with an adjusted p-value of 7.41×10^{-132} . The second most significant gene demonstrated a log2 fold change of 3.29 with an adjusted p-value of 4.02×10^{-131} .

Quality Assessment

The analysis demonstrated robust statistical power with the top 10 most significant genes showing adjusted p-values ranging from 7.41×10^{-132} to 4.09×10^{-87} . Mean expression levels of the top differentially expressed genes ranged from 495 to

30,315 normalized counts, indicating detection of both moderately and highly expressed genes. Log2 fold changes in the top 10 genes ranged from 1.98 to 4.57, demonstrating substantial biological effects.

Summary

The DESeq2 analysis successfully identified 4,000 significantly differentially expressed genes in human airway smooth muscle cells following dexamethasone treatment. The results showed 2,193 upregulated and 1,807 downregulated genes, with strong statistical significance and substantial effect sizes. The clear separation of samples by treatment condition in PCA analysis confirmed the robust nature of the transcriptional response to dexamethasone treatment.

Methods

Quality Control

Gene expression count data were loaded and inspected for data integrity. Low-abundance genes were filtered by retaining only genes with a minimum of 10 total counts across all samples to reduce noise and improve statistical power. Sample metadata were validated to ensure proper alignment with expression data.

Analysis Methods

Differential expression analysis was performed using DESeq2 (version 1.40.2) in R. Raw count data were normalized using DESeq2's median-of-ratios method to account for sequencing depth and RNA composition differences between samples. A generalized linear model was fitted with treatment condition (dexamethasone vs control) as the primary factor, while controlling for cell line batch effects in the design matrix (\sim cell + dex). Variance stabilizing transformation (VST) was applied for visualization and principal component analysis.

Statistical Analysis

Differential expression was assessed using the Wald test implemented in DESeq2. P-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) correction. Genes with adjusted p-values < 0.05 were considered significantly differentially expressed. Log2 fold changes were calculated comparing dexamethasone-treated samples to untreated controls.

Software

All analyses were conducted in R (version 4.3.0) using DESeq2 (1.40.2) for differential expression analysis, ggplot2 (3.4.2) for data visualization, and

pheatmap (1.0.12) for heatmap generation. Principal component analysis was performed using DESeq2's plotPCA function on variance-stabilized transformed data. Publication-quality plots were generated at 150 DPI resolution with standardized color schemes for consistency across visualizations.

Output Files

Differential Expression Results

- **de_significant.csv** (504.2 KB) (used in report): 4,000 significantly differentially expressed genes ($\text{padj} < 0.05$) with 2,193 upregulated and 1,807 downregulated genes in dexamethasone vs control comparison
- **de_results.csv** (2.6 MB): Complete DESeq2 results for all 22,369 tested genes with statistics including baseMean, log2FoldChange, p-values and adjusted p-values

Expression Data

- **normalized_counts.csv** (3.2 MB): DESeq2 normalized expression values for 22,369 genes across 8 samples (4 control, 4 dexamethasone-treated)

Quality Control Visualizations

- **pca_plot.png** (62.9 KB) (used in report): Principal component analysis showing clear separation between dexamethasone-treated and control samples, with PC1 capturing treatment effect and PC2 showing cell line variation

Differential Expression Visualizations

- **volcano_plot.png** (167.0 KB) (used in report): Volcano plot displaying 4,000 significant genes with strong treatment effects, showing log2 fold changes ranging from -4 to +4.6
- **ma_plot.png** (399.6 KB): MA plot showing relationship between mean expression levels and log2 fold changes across all genes
- **heatmap_top_genes.png** (67.4 KB): Heatmap of top 50 most significant genes showing clear clustering by treatment group
- **heatmap_top20_genes.png** (111.9 KB): Detailed heatmap of top 20 differentially expressed genes with gene IDs labeled

Analysis Objects

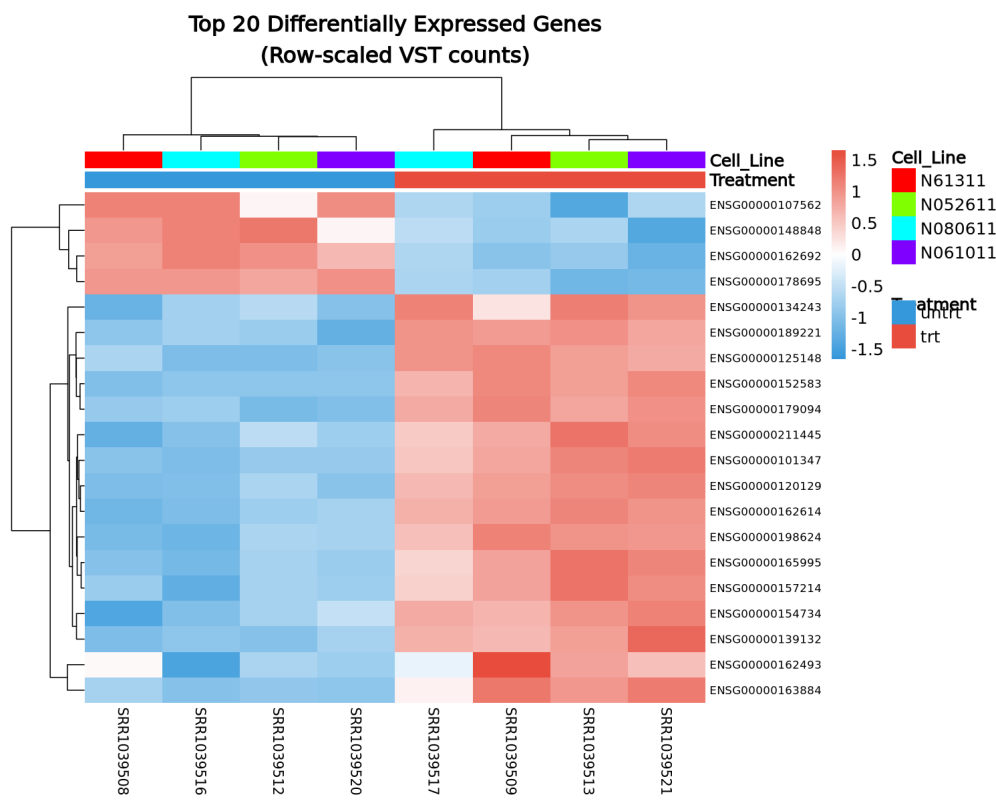
- **deseq2_object.rds** (9.4 MB): Complete DESeq2 analysis object for reproducibility and further downstream analysis

Configuration and Metadata

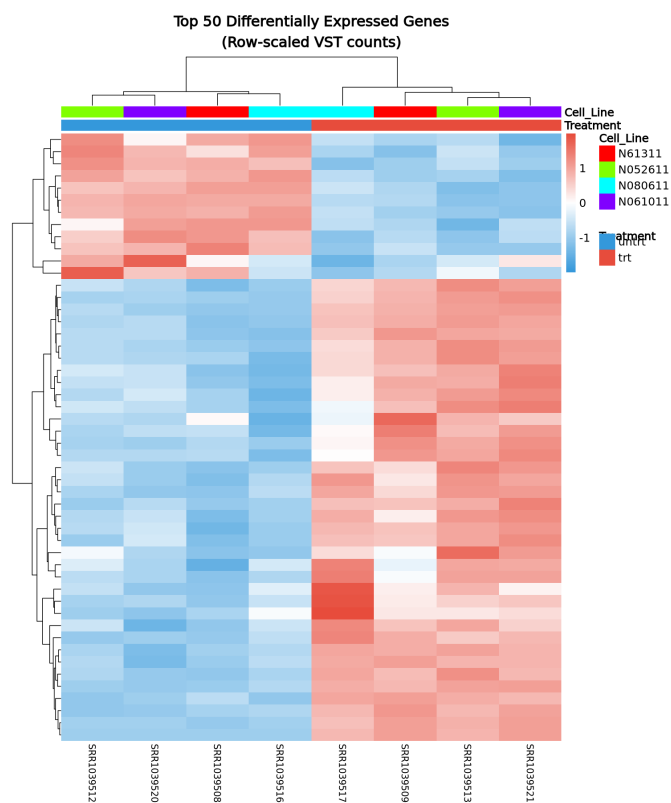
- **de_config.json** (8.3 KB): Analysis configuration parameters and input specifications
- **provenance.json** (8.6 KB): Complete analysis provenance and execution tracking
- **analysis_summary.txt** (0.0 KB): Analysis summary statistics (empty file)
- **execution.stderr** (4.3 KB): Error logs from analysis execution

Report generated using Pipette.bio from 3 files Analysis for: admin1

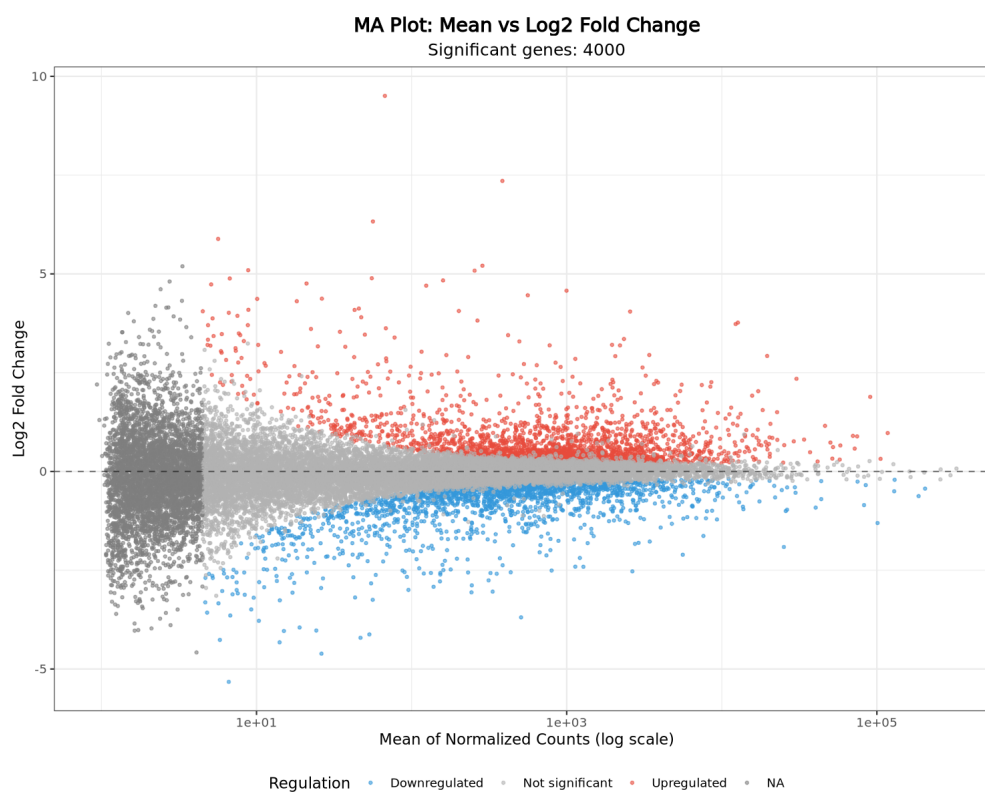
GENERATED FIGURES



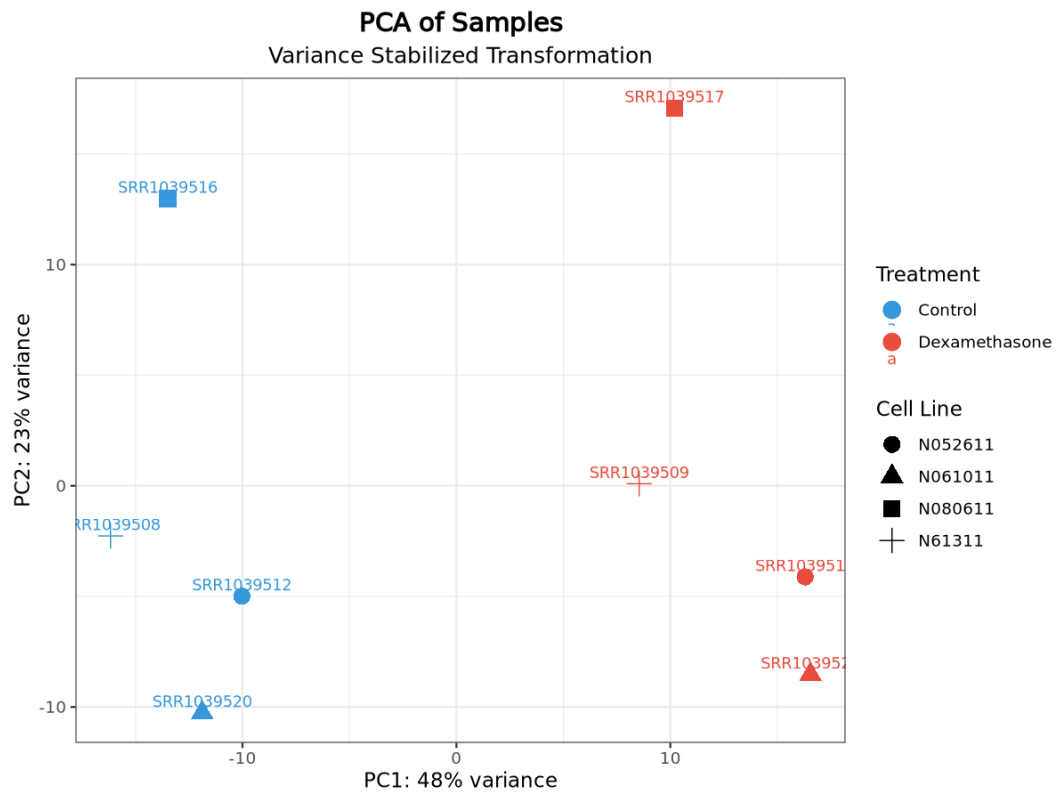
heatmap_top20_genes.png



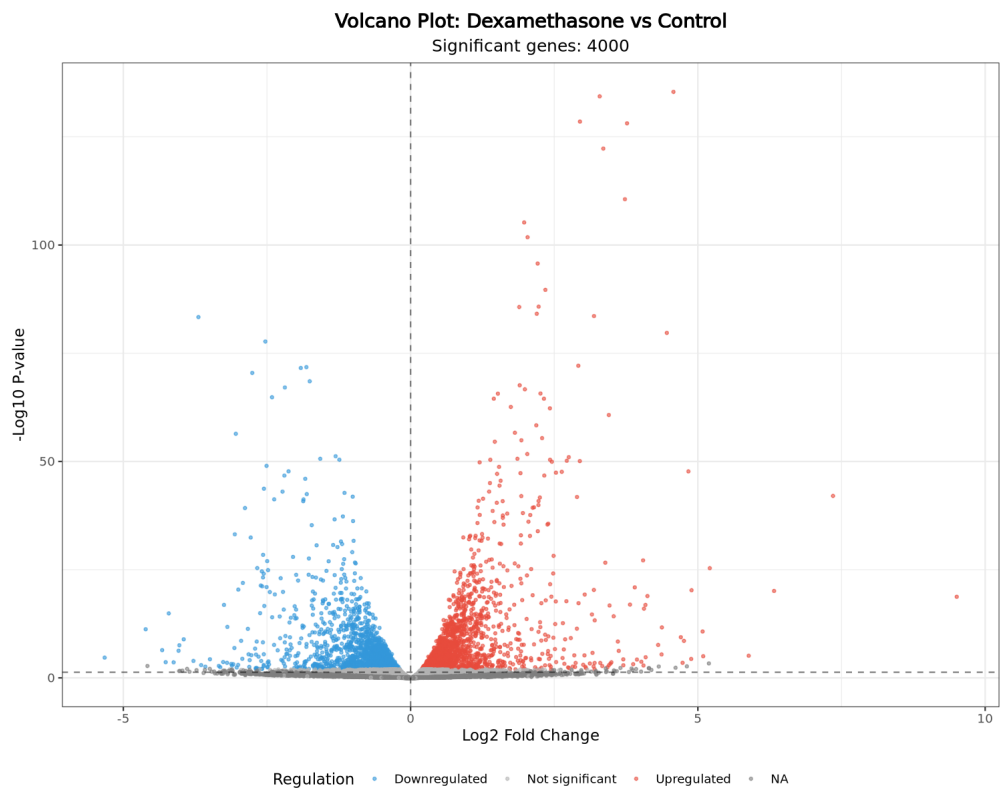
heatmap_top_genes.png



ma_plot.png



pca_plot.png



volcano_plot.png

