**HeraNorm: Endogenous Control Genes and Differentially Expressed Genes Analysis Tool Document**

**Overview**  
HeraNorm is an R Shiny-based application designed to systematically identify and validate endogenous control genes (ECs) and differentially expressed genes (DEGs) from NGS datasets, ensuring their reliable translation into reference genes for qPCR/ddPCR experiments.

**Features**

1. **Data Input and Preprocessing:**

* RNA-Seq or miRNA-Seq count matrices (CSV format)
* Sample metadata information (TXT format).
* Performs normalization and differential expression analysis using a DESeq2 wrapper.

1. **Endogenous Control Genes and DEG Screening:**

* Screens ECs and DEGs based on expression stability and differential expression.
* Allows customization of screening parameters (e.g., p-value threshold, baseMean value, log2FoldChange threshold).

1. **Visualization:**

* Visualizes expression patterns of ECs and DEGs using boxplots.
* Simulates qPCR/ddPCR results by normalizing target genes with user-selected ECs.
* Visualizes DEGs by heatmap and volcano plot.

1. **Data Export:**

* Exports lists of ECs and DEGs as CSV files.
* Exports gene expression plots and In Silico normalization plots as PDF files
* Exports DEG heatmap and volcano plot.

**Prerequisites:**

DEseq2 ≥1.42.1

ggplot2 ≥3.5.1

shiny ≥1.1.4

dplyr ≥1.10.0

pheatmap ≥ 1.0.12

plotly ≥ 4.10.4

RColorBrewer ≥ 1.1.3

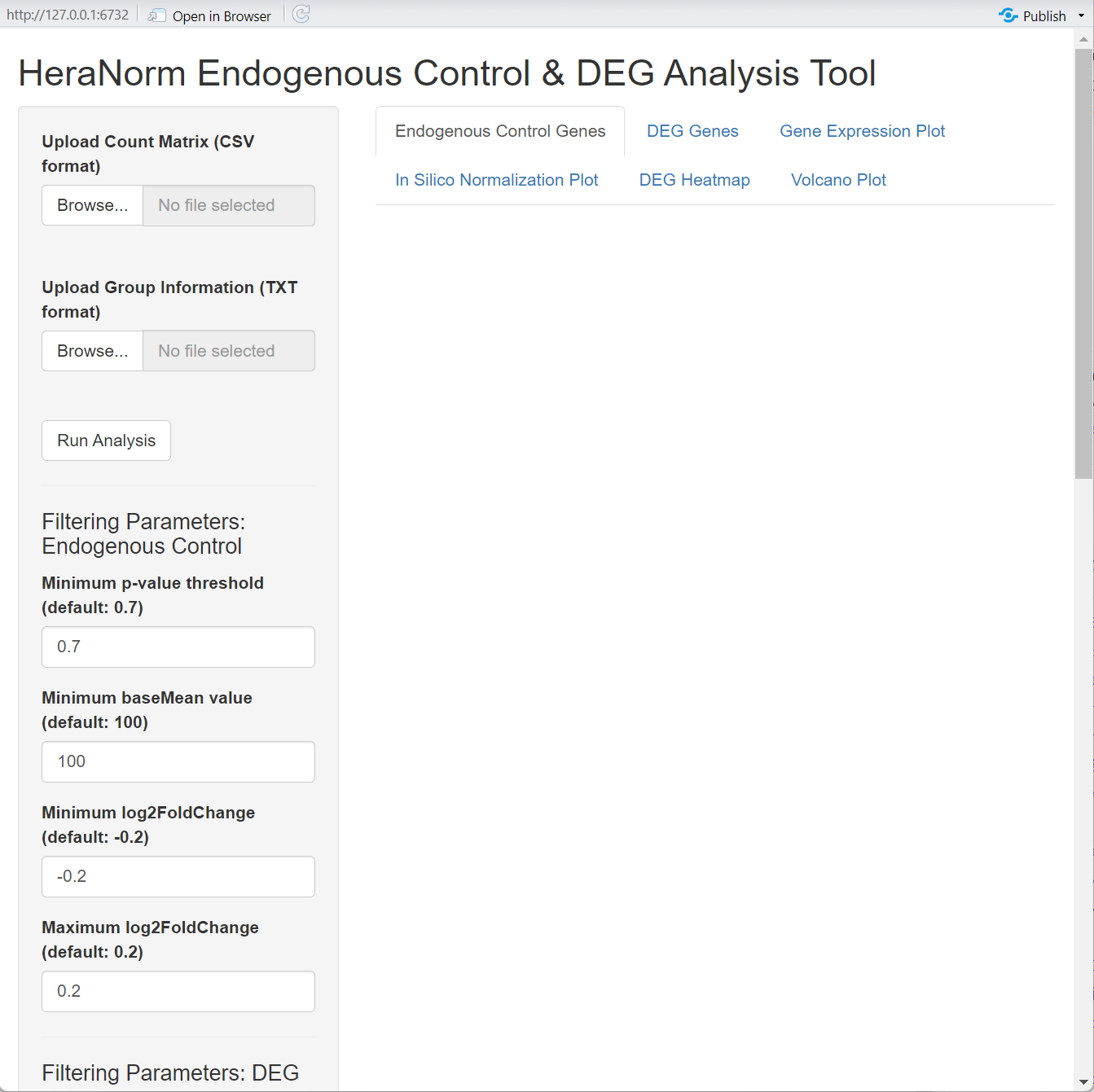
**Usage Instructions**

1. **Launch the Application:**  
   Download the HeraNorm.r app, and save it in your directory. Run the following code in the R environment to start the HeraNorm application:

R

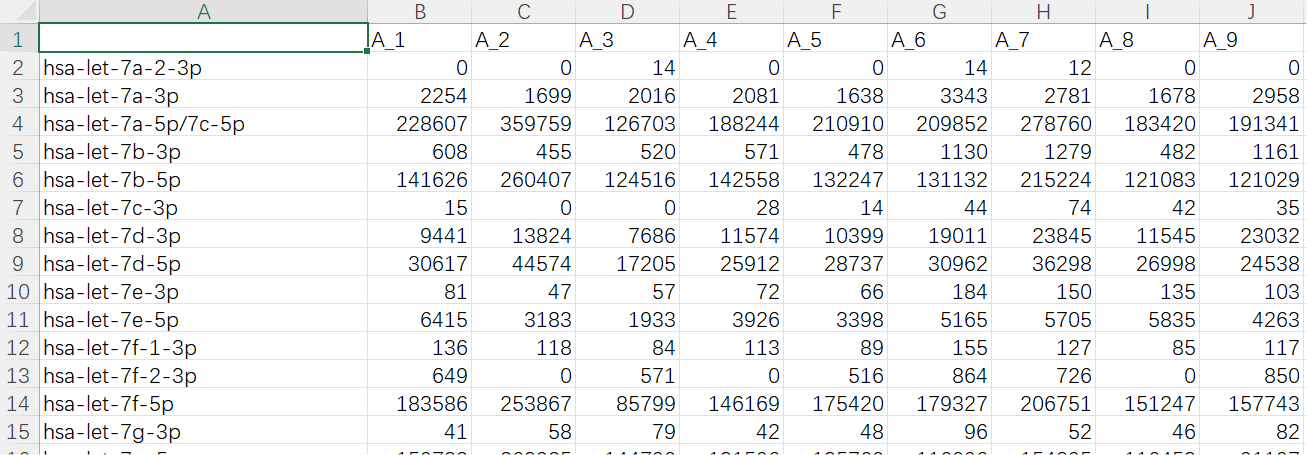
library(shiny)

runApp("path/to/your/HeraNorm/app")

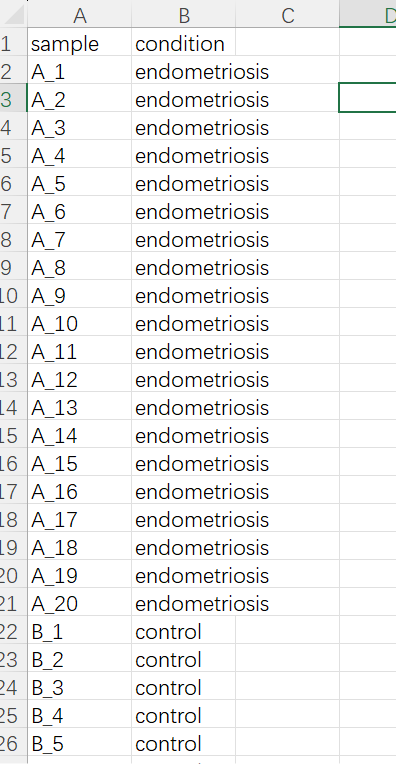
The GUI of the tool: 

1. **Data Input:**

* Upload the RNA-Seq or miRNA-Seq count matrix (CSV format) and sample grouping information (TXT format).
* miRNA-Seq count matrix csv file:



* sample group information:



* Ensure the sample order in the grouping file matches the column order in the count matrix.

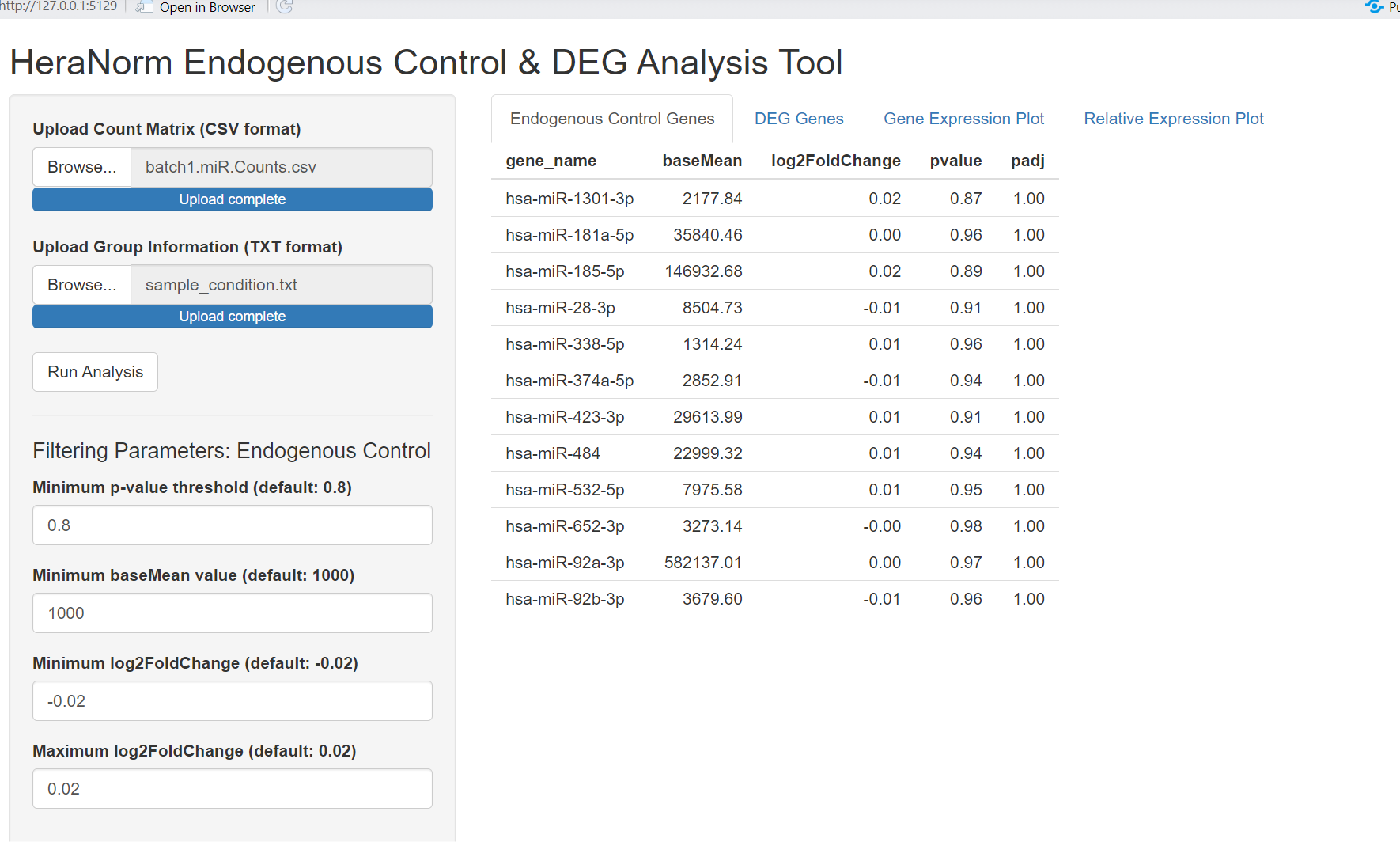
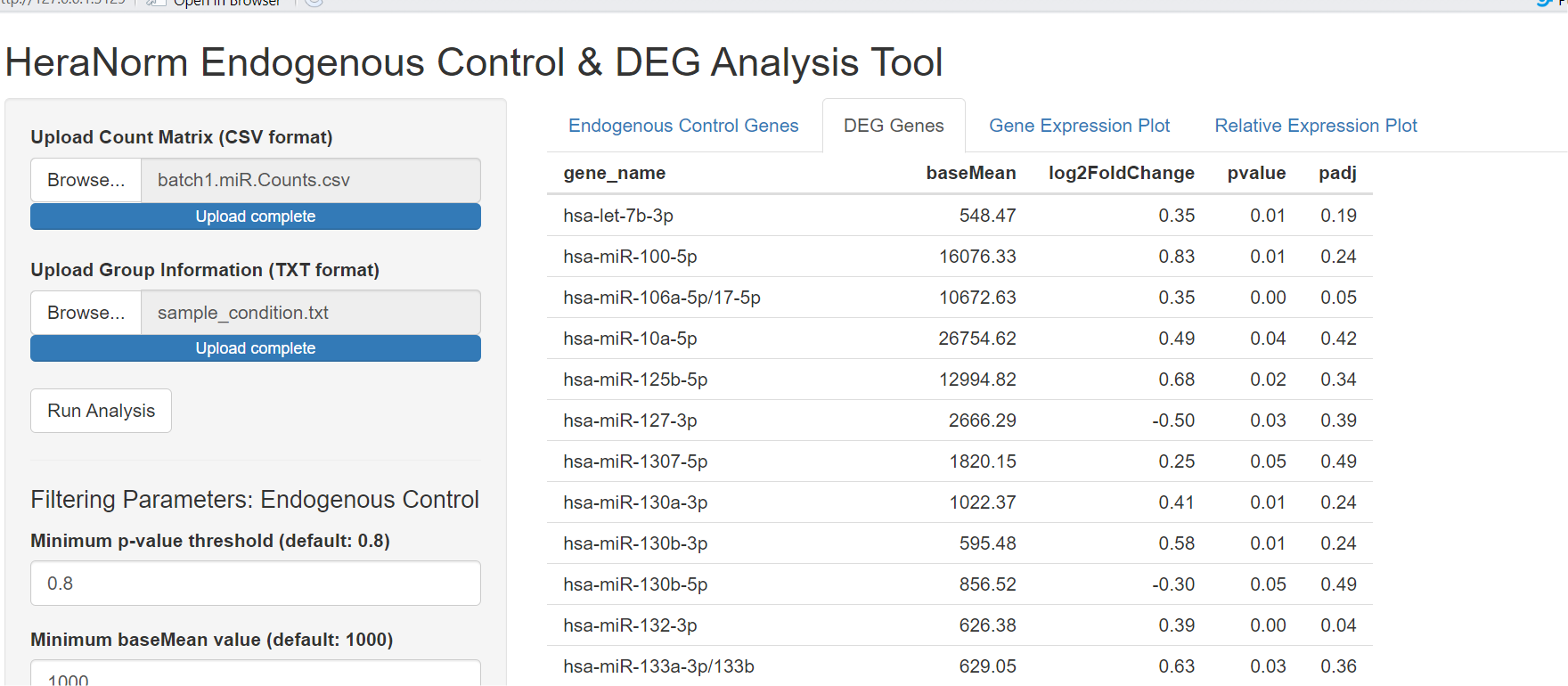
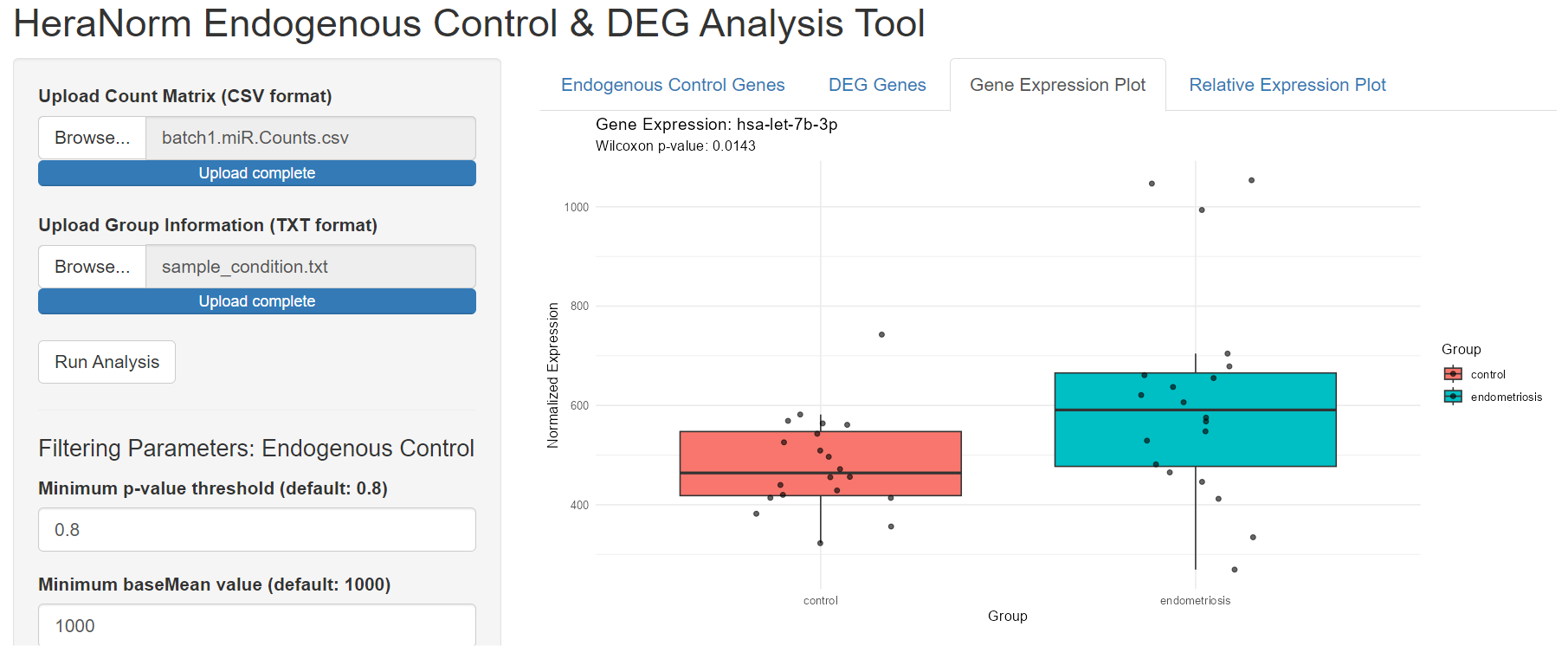
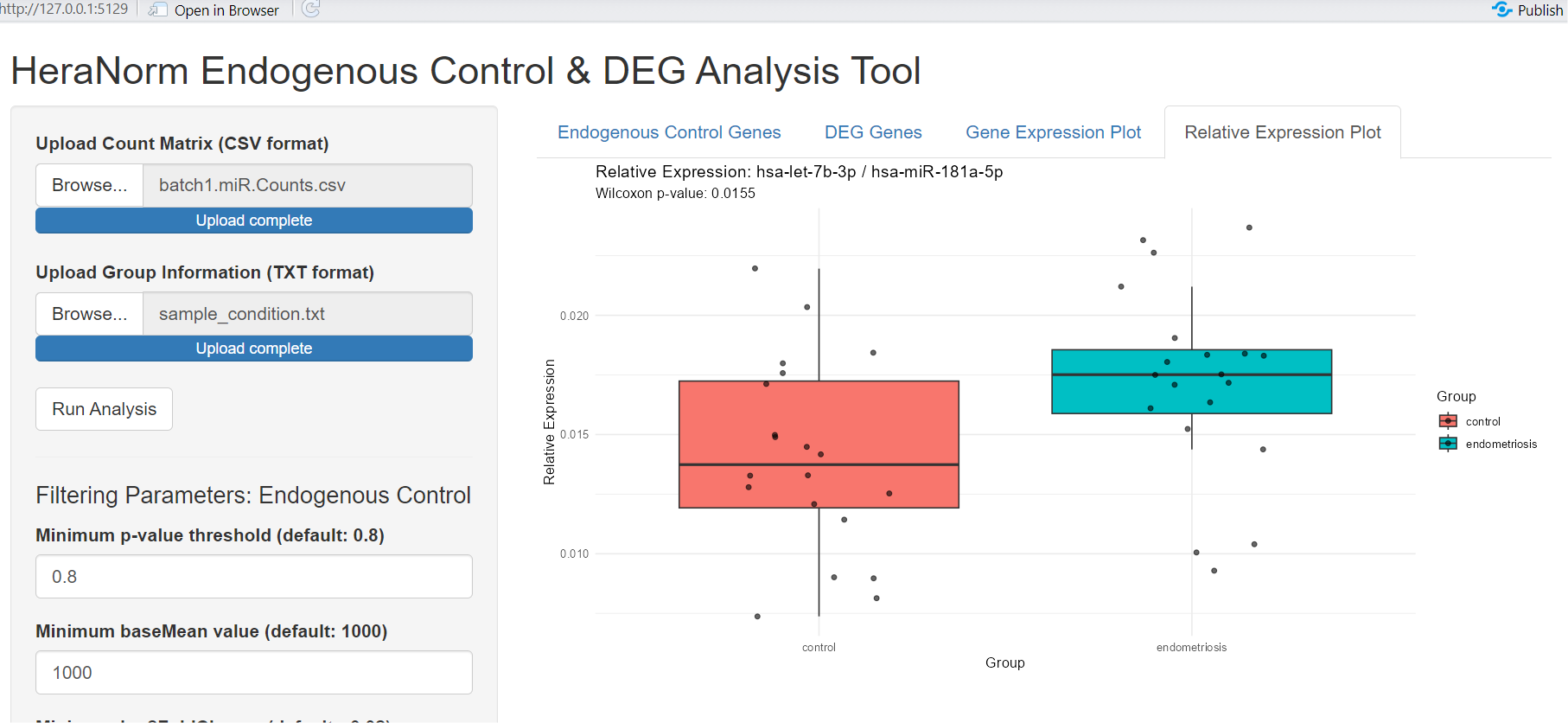
1. **Parameter Settings:**

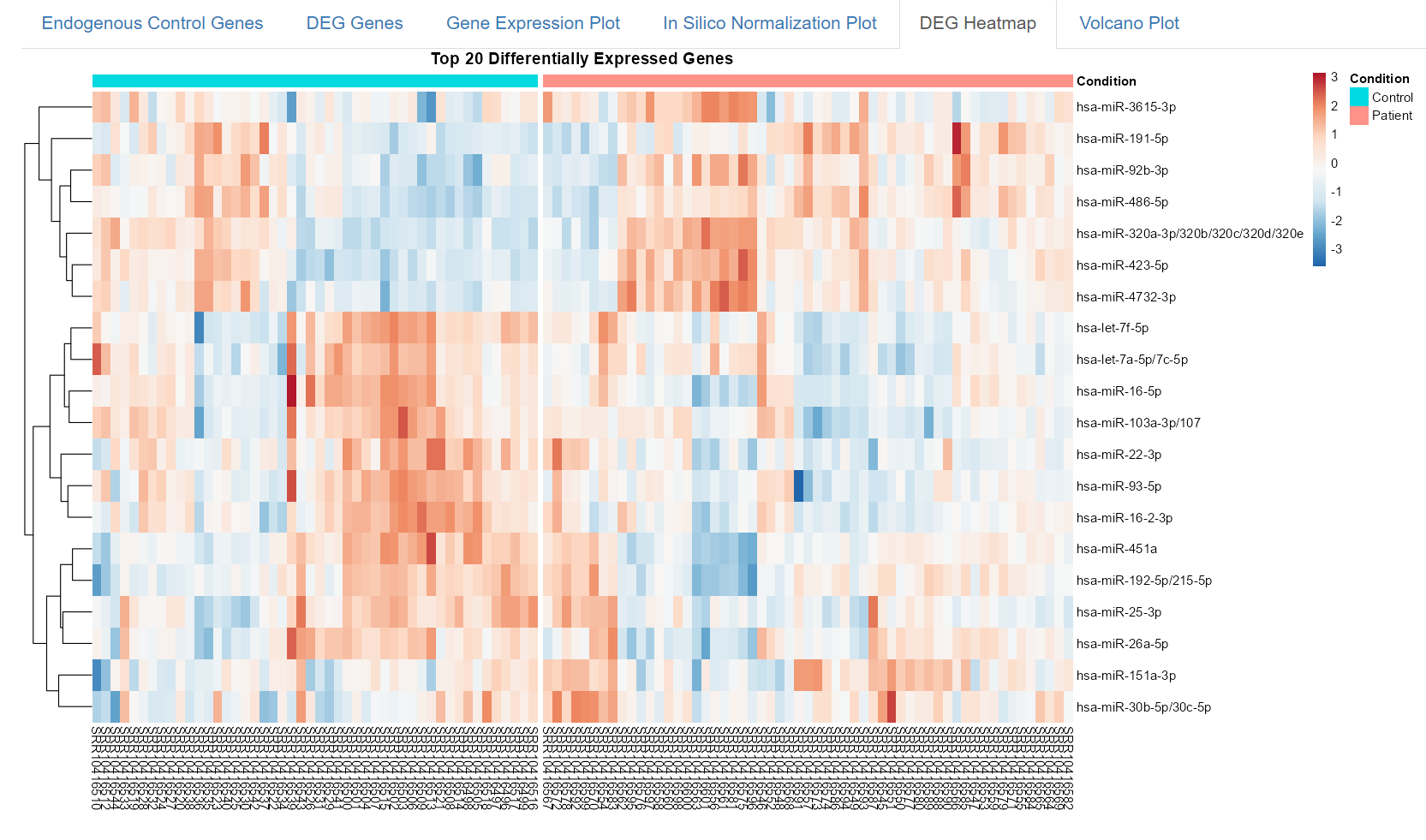
* Set screening parameters for ECs and DEGs (e.g., p-value, baseMean, log2FoldChange).
* Configure visualization parameters (e.g., gene type, gene names).
* Heatmap settings:Number of DEGs to Display(default:50), Color Scheme:three types of color scheme
* Volcano plot settings:Show Gene Labels or not;

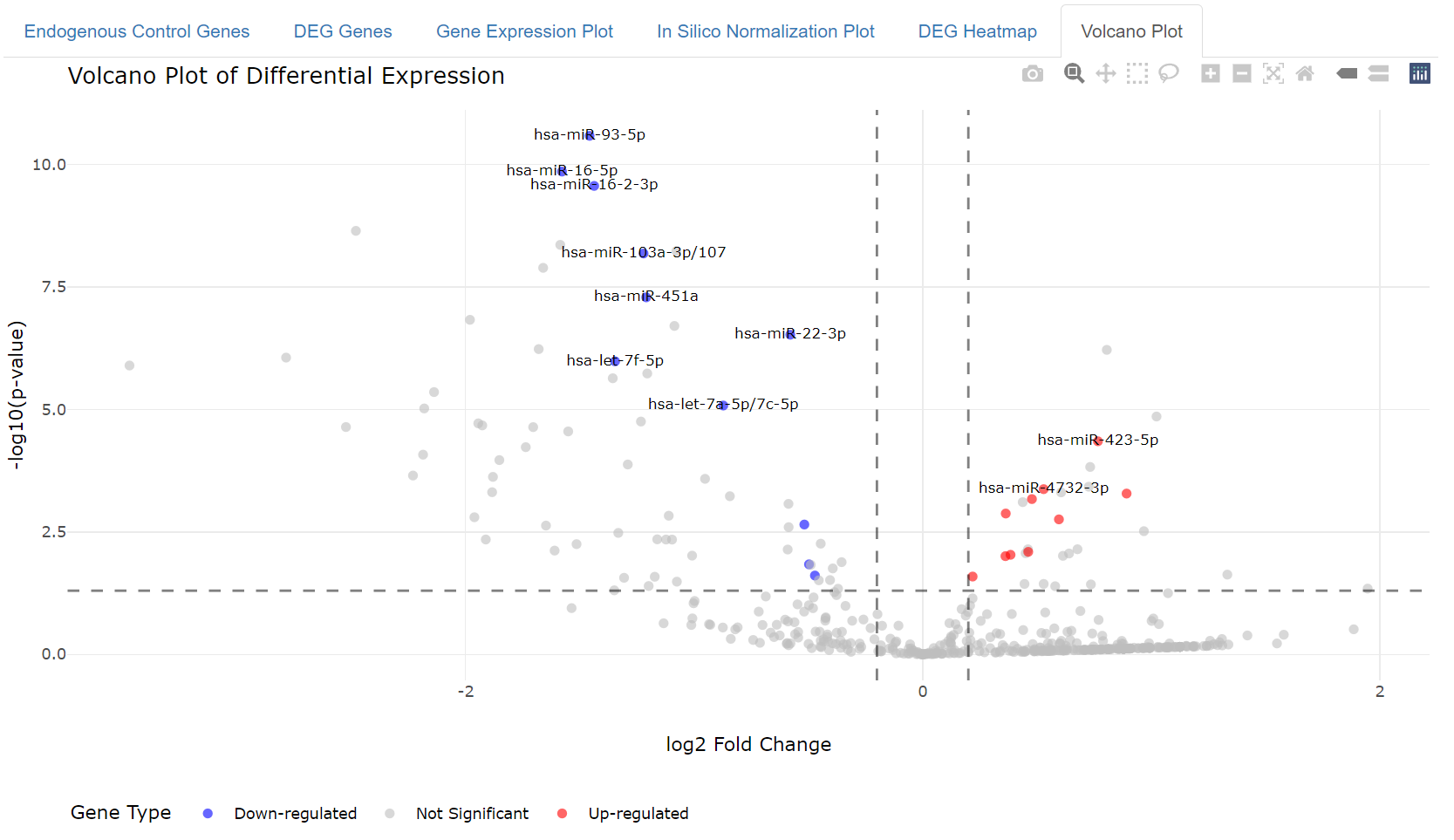
1. **Run Analysis:**  
   Click the **“Run Analysis”** button to initiate the analysis.
2. **View Results:**

* Navigate to the “Endogenous Control Genes” and “DEG Genes” tabs to review filtered EC and DEG lists.
* Use the “Gene Expression Plot” and “In Silico normalization Plot” tabs to visualize expression patterns.
* DEG heatmap plot
* DEG volcano plot
* Export tables and plots using the “Download” button.

**Example**  
Assuming an uploaded miRNA-Seq dataset for endometriosis with predefined screening parameters, you can:

* View identified ECs (e.g., miR-1301-3p and miR-181a-5p) under the **“Endogenous Control Genes”** tab.
* 
* View identified DEGs (e.g., let-7b-3p and miR-100-5p) under the **“DEG Genes”** tab.
* 
* Select let-7b-3p in the **“Gene Expression Plot”** tab to observe its expression across groups.
* 
* Select miR-181a-5p (EC) and let-7b-3p (DEG or Biomarker) in the **“**In Silico normalization Plot**”** tab to compare their relative expression.
* 
* DEG heatmap plot:



* DEG volcano plot
* 
* Export EC/DEG list tables and plots using the “Download” button.

**Notes**

* Verify the correct file formats and sample order in the grouping file before uploading.
* Adjust screening parameters based on your data and research objectives.
* If parameters have been adjusted after the initial run, please rerun the pipeline by clicking ‘Run Analysis’ again.
* Visualization tools aid in interpreting gene expression trends and differential expression.

**Summary**  
HeraNorm is a tool for efficiently identifying and *in silico* validating endogenous control genes and differentially expressed genes from NGS datasets. It enhances the reliability and reproducibility of qPCR/ddPCR experiments, facilitating the translation of genomic research into clinical diagnostic tools.

**Contact**  
For questions or issues, contact us at:  
[wing.h.wong@heranova.com](mailto:wing.h.wong@heranova.com)

[allen.hu@yizhen.com](mailto:allen.hu@yizhen.com)