

# Overall of ViraVate Functionality

- Input:

- Gene Expression Matrix
- Control vs. Disease Key
- Team Identified Genes of Interest
- Optional: Additional User Defined Genes of Interest



- Output:

- Genes associated with viral infection that differ between the Control vs. Disease Groups

## Step 1: Identify Viral Genes of Interest

Database Method

Manual Method



## Step 2: Build User Interface

## Step 1: Identify Viral Genes of Interest

**Database Method:** Identify Known Genes of Interest From Database

Identify top 10% +/- significant genes from Harmonizome

**Manual Method:** Identify Genetic Differences between Infected and Un-infected Populations

SRA db  $\leftarrow$  query[virus]

Raw Counts from RNA-seq data

edgeR to perform Differential Expression

Identify significant DE genes

Identify top 10% +/- significant genes

RSV, sepsis, HSV1, HCMV,  
hepatitis ,C,E,  
plasmodium\_falciparum

Pooled Viral Genes of Interest

## Step 2: Build User Interface

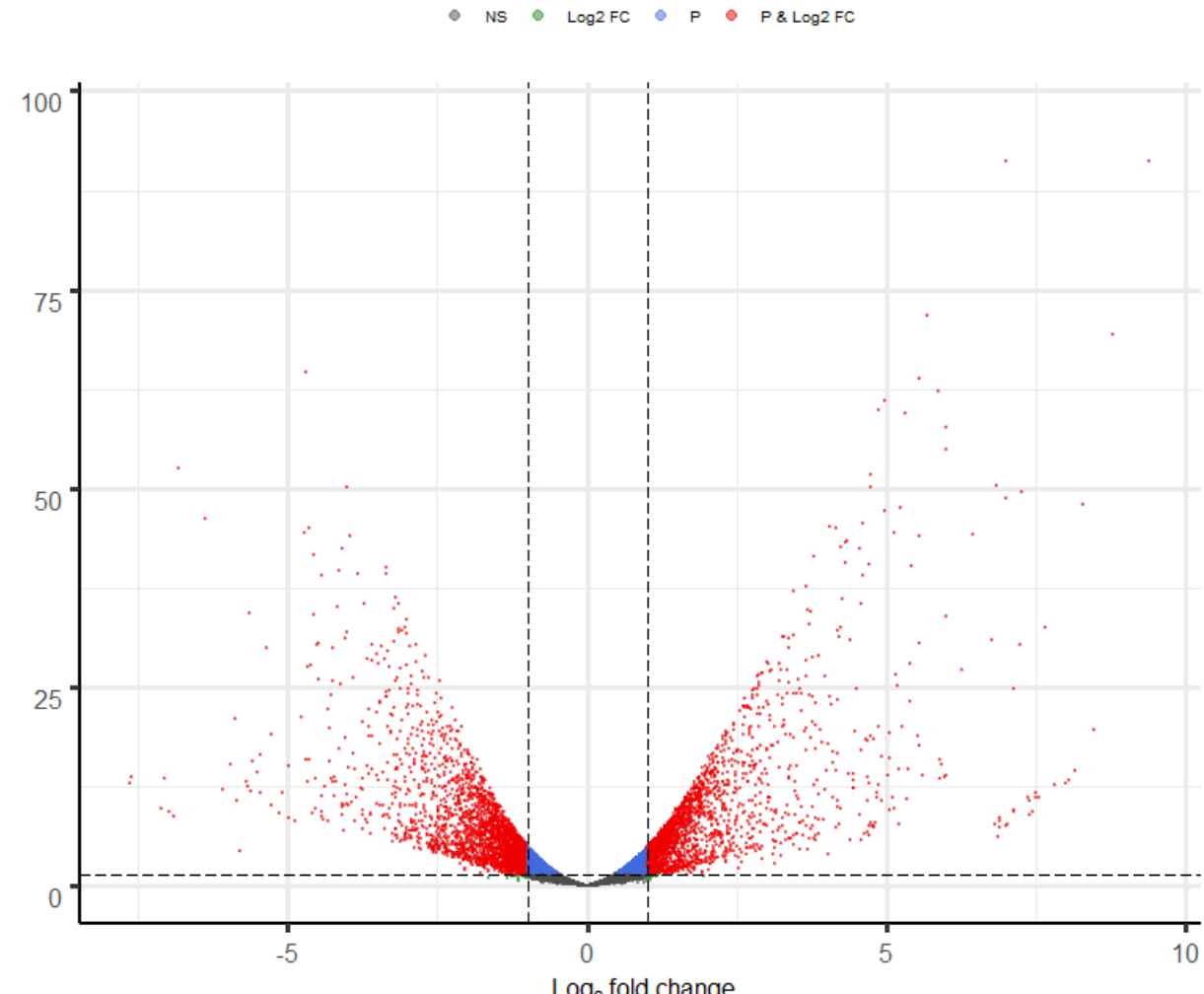
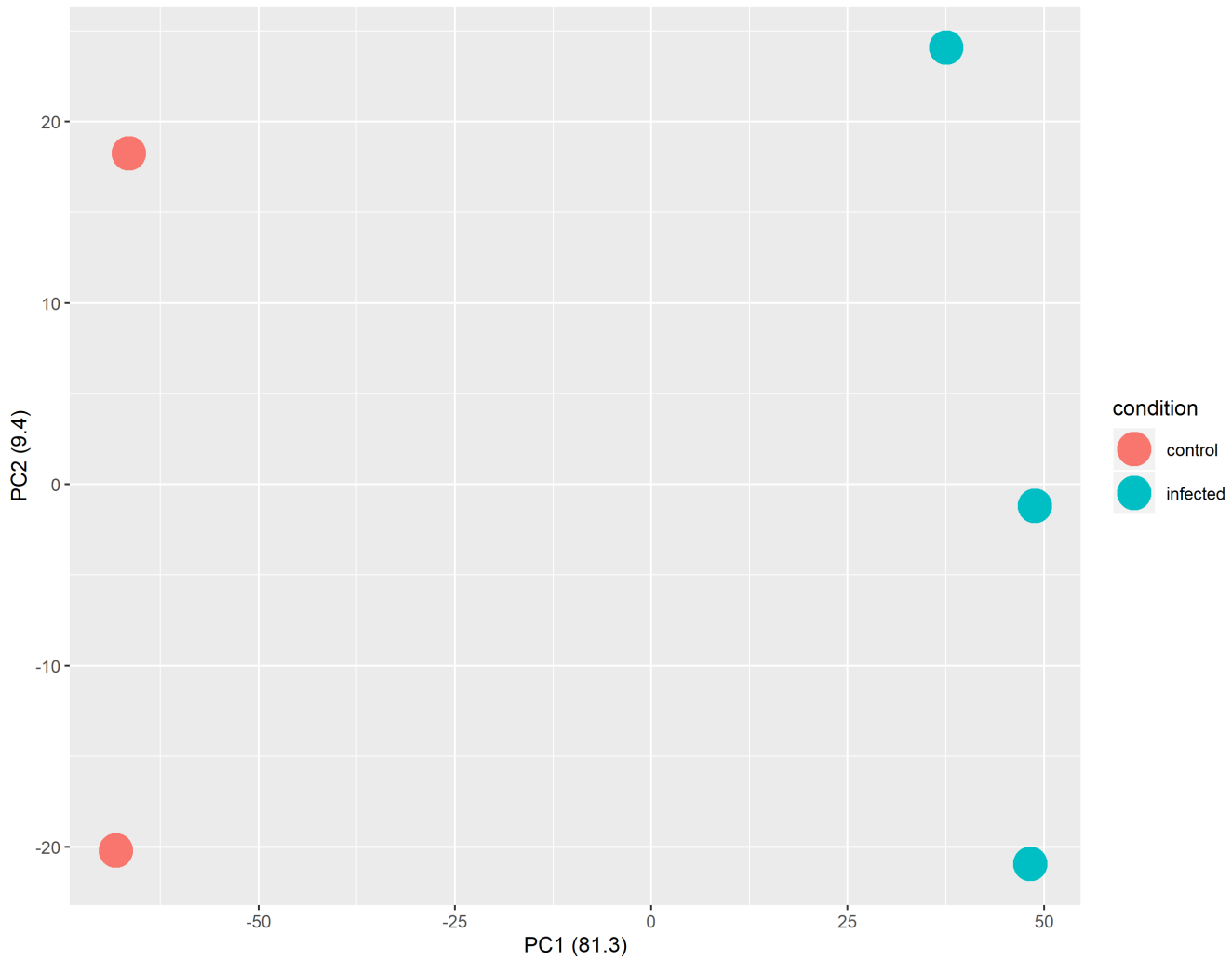
- Algorithm to identify which of the Viral Genes of Interest are different between a disease and control gene expression

Create wrapper  
function for GSECA to  
identify significant  
viruses from database

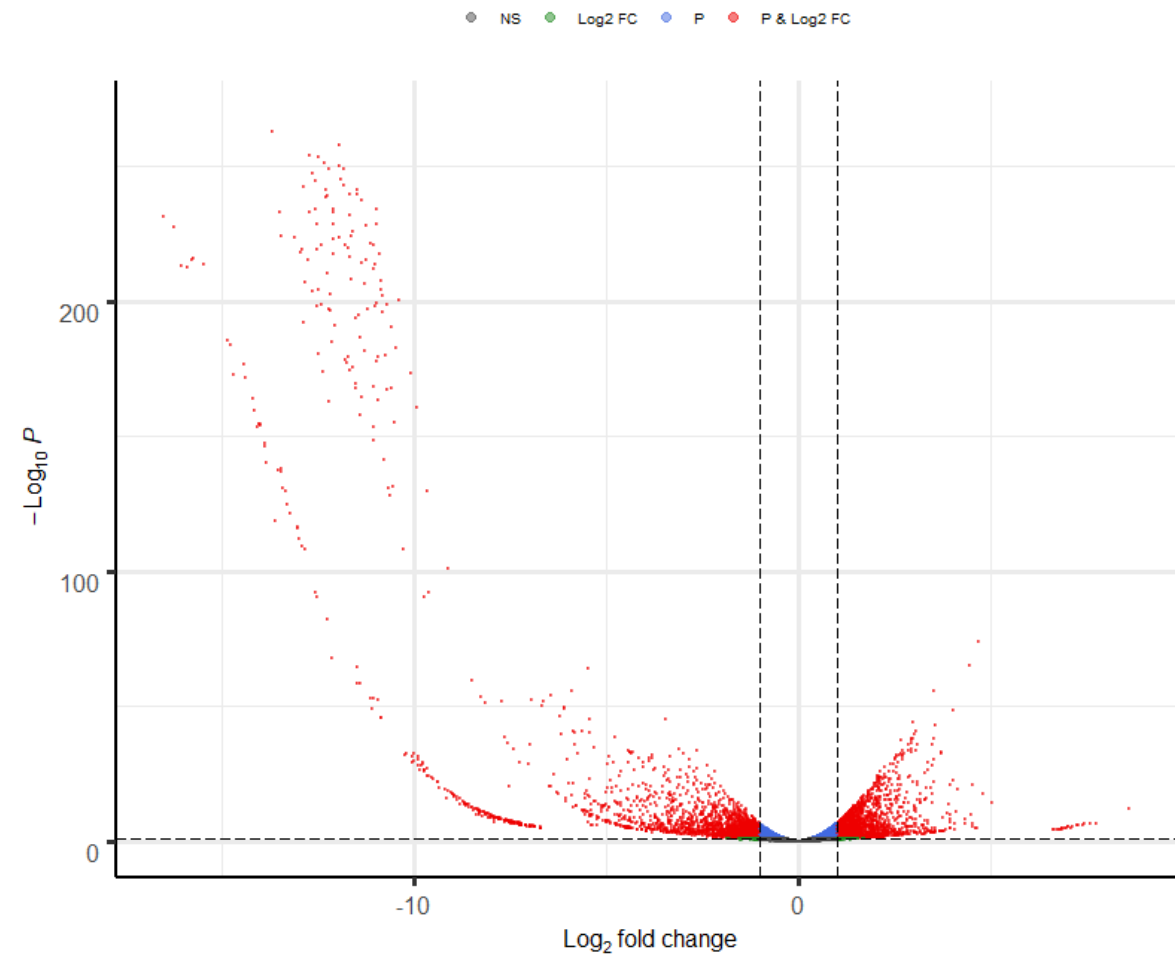
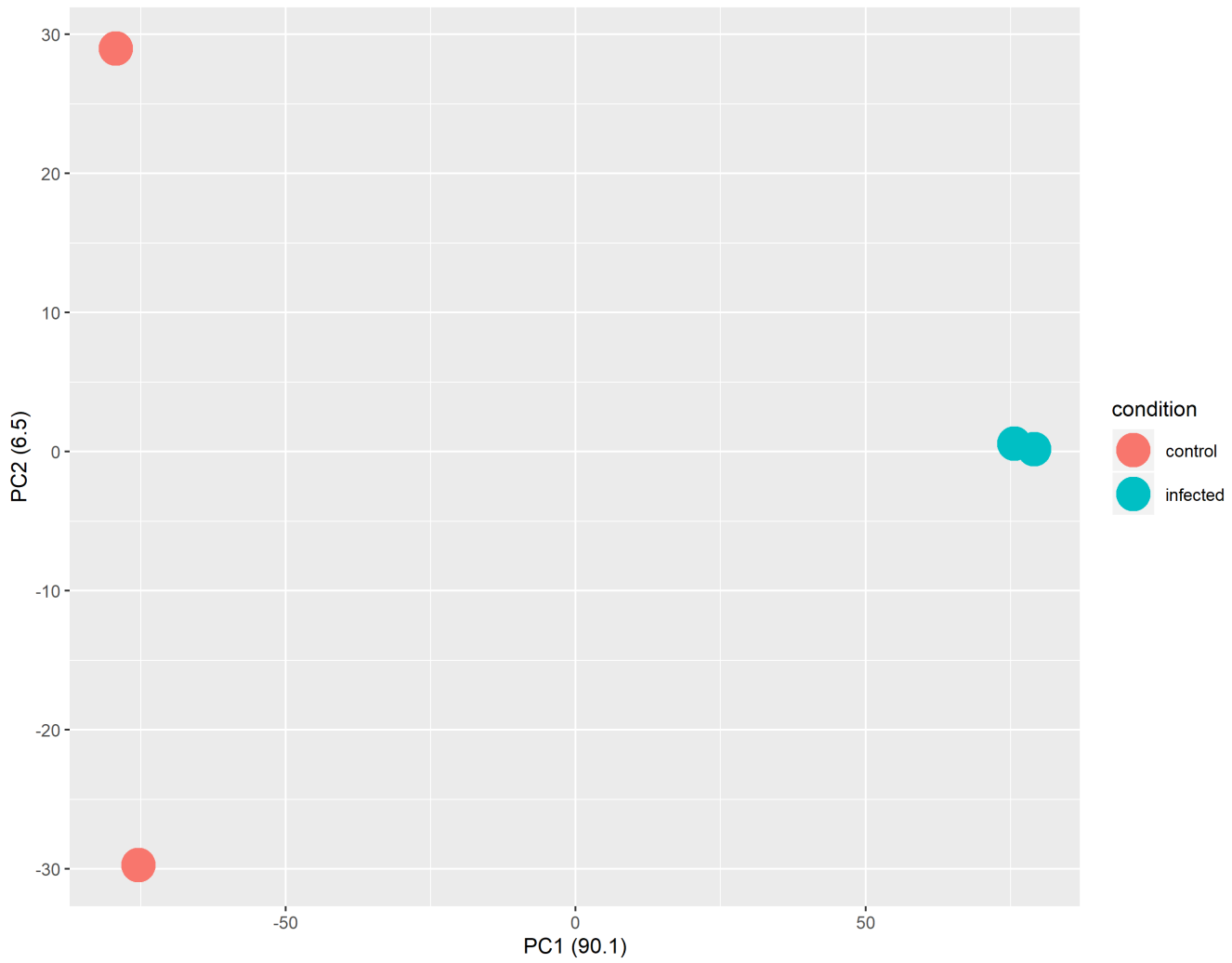


Formal GSECA output  
to clearly show  
significant related  
viruses

# Wild Type Control



# siNT



# siP2Y2

