

assignment2_GeneExpression_DrugResponse_RScript_MeenaEaswaran.R

meena

2025-01-06

```
#set working directory  
#setwd()
```

```
#note the system time  
Sys.time()
```

```
## [1] "2025-01-06 15:30:30 PST"
```

```
#install GEOquery package if required  
#if (!require("BiocManager", quietly = TRUE))  
  #install.packages("BiocManager")  
  
#BiocManager::install("GEOquery", force = TRUE)  
  
#load GEOquery library  
library(GEOquery)
```

```
## Loading required package: Biobase
```

```
## Loading required package: BiocGenerics
```

```
##  
## Attaching package: 'BiocGenerics'
```

```
## The following objects are masked from 'package:stats':  
##  
## IQR, mad, sd, var, xtabs
```

```
## The following objects are masked from 'package:base':  
##  
## anyDuplicated, aperm, append, as.data.frame, basename, cbind,  
## colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,  
## get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,  
## match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,  
## Position, rank, rbind, Reduce, rownames, sapply, setdiff, table,  
## tapply, union, unique, unsplit, which.max, which.min
```

```
## Welcome to Bioconductor
##
## Vignettes contain introductory material; view with
## 'browseVignettes()'. To cite Bioconductor, see
## 'citation("Biobase)"', and for packages 'citation("pkgname)"'.
```

```
## Setting options('download.file.method.GEOquery'='auto')
```

```
## Setting options('GEOquery.inmemory.gpl'=FALSE)
```

```
# Download and read the DataSet SOFT file
```

```
gds <- getGEO(GEO = "GDS3116", filename = "GDS3116.soft.gz")
#gds <- getGEO("GDS3116", GSEMatrix = FALSE, AnnotGPL = FALSE) #This should work too
```

```
##### Preliminary exploration of file to familiarize with GEO dataset #####
```

```
#Explore gds meta data
```

```
Meta(gds)$description
```

```
## [1] "Analysis of breast cancer tumors following treatment with letrozole for 14 days. The aromatase
## [2] "baseline"
## [3] "letrozole"
## [4] "10"
## [5] "11"
## [6] "12"
## [7] "13"
## [8] "14"
## [9] "15"
## [10] "16"
## [11] "17"
## [12] "18"
## [13] "19"
## [14] "20"
## [15] "21"
## [16] "22"
## [17] "23"
## [18] "24"
## [19] "25"
## [20] "27"
## [21] "28"
## [22] "29"
## [23] "2"
## [24] "30"
## [25] "31"
## [26] "33"
## [27] "34"
## [28] "35"
## [29] "36"
## [30] "37"
## [31] "38"
## [32] "39"
## [33] "40"
```

```
## [34] "41"  
## [35] "42"  
## [36] "44"  
## [37] "45"  
## [38] "46"  
## [39] "47"  
## [40] "48"  
## [41] "49"  
## [42] "4"  
## [43] "50"  
## [44] "51"  
## [45] "52"  
## [46] "53"  
## [47] "55"  
## [48] "56"  
## [49] "57"  
## [50] "58"  
## [51] "59"  
## [52] "5"  
## [53] "60"  
## [54] "63"  
## [55] "64"  
## [56] "65"  
## [57] "68"  
## [58] "69"  
## [59] "7"  
## [60] "8"  
## [61] "9"
```

```
Meta(gds)$platform
```

```
## [1] "GPL96"
```

```
Meta(gds)$sample_count
```

```
## [1] "116"
```

```
Meta(gds)$sample_organism
```

```
## [1] "Homo sapiens"
```

```
Meta(gds)$sample_type
```

```
## [1] "RNA"
```

```
Meta(gds)$title
```

```
## [1] "Letrozole effect on breast cancer tumors"
```

```
#Explore sample information associated to each phenotype
gds@dataTable@columns[["description"]]
```

```
## [1] "Value for GSM125123: 10A;pretreatment;female;breast tumor; responder; src: 10A;pretreatment;f
## [2] "Value for GSM125125: 11A;pretreatment;female;breast tumor; nonresponder; src: 11A;pretreatmen
## [3] "Value for GSM125127: 12A;pretreatment;female;breast tumor; responder; src: 12A;pretreatment;f
## [4] "Value for GSM125129: 13A;pretreatment;female;breast tumor; responder; src: 13A;pretreatment;f
## [5] "Value for GSM125131: 14A;pretreatment;female;breast tumor; responder; src: 14A;pretreatment;f
## [6] "Value for GSM125133: 15A;pretreatment;female;breast tumor; responder; src: 15A;pretreatment;f
## [7] "Value for GSM125135: 16A;pretreatment;female;breast tumor; not assessable; src: 16A;pretreatm
## [8] "Value for GSM125137: 17A;pretreatment;female;breast tumor; responder; src: 17A;pretreatment;f
## [9] "Value for GSM125139: 18A;pretreatment;female;breast tumor; responder; src: 18A;pretreatment;f
## [10] "Value for GSM125141: 19A;pretreatment;female;breast tumor; responder; src: 19A;pretreatment;f
## [11] "Value for GSM125143: 20A;pretreatment;female;breast tumor; responder; src: 20A;pretreatment;f
## [12] "Value for GSM125145: 21A;pretreatment;female;breast tumor; responder; src: 21A;pretreatment;f
## [13] "Value for GSM125147: 22A;pretreatment;female;breast tumor; responder; src: 22A;pretreatment;f
## [14] "Value for GSM125149: 23A;pretreatment;female;breast tumor; responder; src: 23A;pretreatment;f
## [15] "Value for GSM125151: 24A;pretreatment;female;breast tumor; not assessable; src: 24A;pretreatm
## [16] "Value for GSM125153: 25A;pretreatment;female;breast tumor; responder; src: 25A;pretreatment;f
## [17] "Value for GSM125155: 27A;pretreatment;female;breast tumor; nonresponder; src: 27A;pretreatmen
## [18] "Value for GSM125157: 28A;pretreatment;female;breast tumor; responder; src: 28A;pretreatment;f
## [19] "Value for GSM125159: 29A;pretreatment;female;breast tumor; responder; src: 29A;pretreatment;f
## [20] "Value for GSM125161: 2A;pretreatment;female;breast tumor; responder; src: 2A;pretreatment;fema
## [21] "Value for GSM125163: 30A;pretreatment;female;breast tumor; nonresponder; src: 30A;pretreatmen
## [22] "Value for GSM125165: 31A;pretreatment;female;breast tumor; responder; src: 31A;pretreatment;f
## [23] "Value for GSM125167: 33A;pretreatment;female;breast tumor; nonresponder; src: 33A;pretreatmen
## [24] "Value for GSM125169: 34A;pretreatment;female;breast tumor; responder; src: 34A;pretreatment;f
## [25] "Value for GSM125171: 35A;pretreatment;female;breast tumor; nonresponder; src: 35A;pretreatmen
## [26] "Value for GSM125173: 36A;pretreatment;female;breast tumor; nonresponder; src: 36A;pretreatmen
## [27] "Value for GSM125175: 37A;pretreatment;female;breast tumor; not assessable; src: 37A;pretreatm
## [28] "Value for GSM125177: 38A;pretreatment;female;breast tumor; nonresponder; src: 38A;pretreatmen
## [29] "Value for GSM125179: 39A;pretreatment;female;breast tumor; responder; src: 39A;pretreatment;f
## [30] "Value for GSM125181: 40A;pretreatment;female;breast tumor; responder; src: 40A;pretreatment;f
## [31] "Value for GSM125183: 41A;pretreatment;female;breast tumor; responder; src: 41A;pretreatment;f
## [32] "Value for GSM125185: 42A;pretreatment;female;breast tumor; nonresponder; src: 42A;pretreatmen
## [33] "Value for GSM125187: 44A;pretreatment;female;breast tumor; nonresponder; src: 44A;pretreatmen
## [34] "Value for GSM125189: 45A;pretreatment;female;breast tumor; responder; src: 45A;pretreatment;f
## [35] "Value for GSM125191: 46A;pretreatment;female;breast tumor; not assessable; src: 46A;pretreatm
## [36] "Value for GSM125193: 47A;pretreatment;female;breast tumor; responder; src: 47A;pretreatment;f
## [37] "Value for GSM125195: 48A;pretreatment;female;breast tumor; responder; src: 48A;pretreatment;f
## [38] "Value for GSM125197: 49A;pretreatment;female;breast tumor; responder; src: 49A;pretreatment;f
## [39] "Value for GSM125199: 4A;pretreatment;female;breast tumor; not assessable; src: 4A;pretreatmen
## [40] "Value for GSM125201: 50A;pretreatment;female;breast tumor; responder; src: 50A;pretreatment;f
## [41] "Value for GSM125203: 51A;pretreatment;female;breast tumor; responder; src: 51A;pretreatment;f
## [42] "Value for GSM125205: 52A;pretreatment;female;breast tumor; nonresponder; src: 52A;pretreatmen
## [43] "Value for GSM125207: 53A;pretreatment;female;breast tumor; nonresponder; src: 53A;pretreatmen
## [44] "Value for GSM125209: 55A;pretreatment;female;breast tumor; responder; src: 55A;pretreatment;f
## [45] "Value for GSM125211: 56A;pretreatment;female;breast tumor; not assessable; src: 56A;pretreatm
## [46] "Value for GSM125213: 57A;pretreatment;female;breast tumor; responder; src: 57A;pretreatment;f
## [47] "Value for GSM125215: 58A;pretreatment;female;breast tumor; responder; src: 58A;pretreatment;f
## [48] "Value for GSM125217: 59A;pretreatment;female;breast tumor; nonresponder; src: 59A;pretreatmen
## [49] "Value for GSM125219: 5A;pretreatment;female;breast tumor; responder; src: 5A;pretreatment;fema
## [50] "Value for GSM125221: 60A;pretreatment;female;breast tumor; responder; src: 60A;pretreatment;f
```

[51] "Value for GSM125223: 63A;pretreatment;female;breast tumor; responder; src: 63A;pretreatment;f
 ## [52] "Value for GSM125225: 64A;pretreatment;female;breast tumor; responder; src: 64A;pretreatment;f
 ## [53] "Value for GSM125227: 65A;pretreatment;female;breast tumor; nonresponder; src: 65A;pretreatment
 ## [54] "Value for GSM125229: 68A;pretreatment;female;breast tumor; nonresponder; src: 68A;pretreatment
 ## [55] "Value for GSM125231: 69A;pretreatment;female;breast tumor; responder; src: 69A;pretreatment;f
 ## [56] "Value for GSM125233: 7A;pretreatment;female;breast tumor; responder; src: 7A;pretreatment;fem
 ## [57] "Value for GSM125235: 8A;pretreatment;female;breast tumor; nonresponder; src: 8A;pretreatment;
 ## [58] "Value for GSM125237: 9A;pretreatment;female;breast tumor; responder; src: 9A;pretreatment;fem
 ## [59] "Value for GSM125124: 10B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [60] "Value for GSM125126: 11B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [61] "Value for GSM125128: 12B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [62] "Value for GSM125130: 13B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [63] "Value for GSM125132: 14B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [64] "Value for GSM125134: 15B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [65] "Value for GSM125136: 16B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; not asses
 ## [66] "Value for GSM125138: 17B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [67] "Value for GSM125140: 18B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [68] "Value for GSM125142: 19B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [69] "Value for GSM125144: 20B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [70] "Value for GSM125146: 21B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [71] "Value for GSM125148: 22B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [72] "Value for GSM125150: 23B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [73] "Value for GSM125152: 24B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; not asses
 ## [74] "Value for GSM125154: 25B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [75] "Value for GSM125156: 27B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [76] "Value for GSM125158: 28B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [77] "Value for GSM125160: 29B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [78] "Value for GSM125162: 2B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder;
 ## [79] "Value for GSM125164: 30B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [80] "Value for GSM125166: 31B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [81] "Value for GSM125168: 33B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [82] "Value for GSM125170: 34B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [83] "Value for GSM125172: 35B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [84] "Value for GSM125174: 36B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [85] "Value for GSM125176: 37B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; not asses
 ## [86] "Value for GSM125178: 38B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [87] "Value for GSM125180: 39B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [88] "Value for GSM125182: 40B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [89] "Value for GSM125184: 41B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [90] "Value for GSM125186: 42B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [91] "Value for GSM125188: 44B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [92] "Value for GSM125190: 45B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [93] "Value for GSM125192: 46B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; not asses
 ## [94] "Value for GSM125194: 47B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [95] "Value for GSM125196: 48B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [96] "Value for GSM125198: 49B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [97] "Value for GSM125200: 4B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; not assess
 ## [98] "Value for GSM125202: 50B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [99] "Value for GSM125204: 51B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [100] "Value for GSM125206: 52B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [101] "Value for GSM125208: 53B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonrespon
 ## [102] "Value for GSM125210: 55B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder
 ## [103] "Value for GSM125212: 56B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; not asses
 ## [104] "Value for GSM125214: 57B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder

```
## [105] "Value for GSM125216: 58B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder"
## [106] "Value for GSM125218: 59B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonresponder"
## [107] "Value for GSM125220: 5B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder;"
## [108] "Value for GSM125222: 60B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder"
## [109] "Value for GSM125224: 63B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder"
## [110] "Value for GSM125226: 64B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder"
## [111] "Value for GSM125228: 65B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonresponder"
## [112] "Value for GSM125230: 68B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonresponder"
## [113] "Value for GSM125232: 69B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder"
## [114] "Value for GSM125234: 7B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder;"
## [115] "Value for GSM125236: 8B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; nonresponder"
## [116] "Value for GSM125238: 9B;female;breast tumor;Letrozole, 2.5mg/day,oral, 10-14 days; responder;"
```

```
#Explore probe information if required
#gds@dataTable@table[["ID_REF"]]
```

```
##### Convert GDS to ExpressionSet #####
```

```
#Convert GDS to an ExpressionSet
eset <- GDS2eSet(gds,do.log2=FALSE) #do.log2 as TRUE will log transform
```

```
#View phenotype data in Expressionset as a data frame
phenotype_df <- as.data.frame(pData(eset))
View(phenotype_df)
#This file shows that required information about two relevant phenotypes: "responder" and "nonresponder"
```

```
##### Perform phenotype filtering based on description column #####
```

```
#1: Convert ExpressionSet as a data frame
eset_df <- as.data.frame(eset)
View(eset_df)
#This will have X prefixed to all probes IDs in the data frame columns except for the last 4 columns
```

```
# Retrieve actual IDs and assign them as column names of the data frame
colnames(eset_df) <- gds@dataTable@table[["ID_REF"]]
```

```
# Assign specific names as earlier in the phenotype data to the last four columns
colnames(eset_df)[(ncol(eset_df) - 3):ncol(eset_df)] <- c("sample", "agent", "individual", "description")
View(eset_df) #All Column IDs restored to original
```

```
#2: Filter rows with only "responder" or "non-responder"phenotypes into a new data frame
filtered_df <- eset_df[grepl("responder|nonresponder", eset_df$description) &
!grepl("not assessable", eset_df$description), ]
# This data frame has removed 12 rows or samples which had a phenotype of "non assessable"
View(filtered_df)
```

```
# Confirm if there is "not assessable" phenotype at the end of the filtered data
not_assessable_at_end <- grepl("not assessable$", filtered_df$description)
# Get total sum of the above logical vector
sum_not_assessable_at_end <- sum(not_assessable_at_end)
sum_not_assessable_at_end #This should return 0
```

```
## [1] 0
```

```
##### Data frame cleanup and assigning "responder" or "nonresponder" phenotype annotations to

#Remove unwanted columns with non-numeric data except for "description" before differential expression
# Create a new data frame
new_df <- filtered_df[, !(names(filtered_df) %in% c("sample", "agent", "individual"))]
View(new_df)

#Annotate phenotype information for each row using description column
#Create a function to detect the presence of "responder" or "nonresponder"
detect_phenotype <- function(description) {
  if (grepl("responder", description, ignore.case = TRUE) && !grepl("nonresponder", description, ignore
    return("responder")
  } else if (grepl("nonresponder", description, ignore.case = TRUE)) {
    return("nonresponder")
  } else {
    return(NA) # In case neither is found
  }
}

# Apply the function to each element of the description column
phenotypes <- sapply(new_df[, "description"], detect_phenotype)

# Add a new column with detected phenotype information to the data frame
new_df["phenotype"] = phenotypes
View(new_df)
#The phenotype information in description column should match the "responder" or "nonresponder" status

# Create a new data frame with description column deleted
new_df_1 <- new_df[, !grepl("description", colnames(new_df))]
View(new_df_1) #This should have phenotype as the last column and description column deleted

#####Differential expression using t -test with equal variance and multiple hypothesis test with Bonfer

#install stats package if required
#install.packages("stats")

# Load stats library if required
#library(stats)

# Check for null values
if (any(is.na(new_df_1))) {
  # Replace null values with zeros
  new_df_1[is.na(new_df_1)] <- 0
}

# Split the data into "responder" and "nonresponder" groups to get an idea about the sample sizes across
responder_data <- new_df_1[new_df_1$phenotype == "responder", ]
View(responder_data)
print(nrow(responder_data)) #This should yield "responder" sample size
```

```
## [1] 74
```

```

nonresponder_data <- new_df_1[new_df_1$phenotype == "nonresponder", ]
View(nonresponder_data)
print(nrow(nonresponder_data)) #This should yield "nonresponder" sample size

```

```
## [1] 30
```

```

#The sample sizes between groups are unequal
#Welch t-test maybe more robust in this condition
#However, perform subsequent t-tests assuming equal variance as instructed

#Perform a t-test comparing columns(probes) between two groups defined by the phenotype variable
t_test_results <- apply(new_df_1[, -ncol(new_df_1)], 2, function(probe) {
  t_test <- t.test(probe ~ new_df_1$phenotype, var.equal = TRUE)
  return(c(p_value = t_test$p.value, t_stat = t_test$statistic))
})
View(t_test_results)

# Transpose t_test_results and convert into a data frame
t_test_results_df <- as.data.frame(t(t_test_results))

# Compute Bonferroni-corrected p-values
t_test_results_df$Bonferroni_adjusted_p_value <- p.adjust(t_test_results_df$p_value, method = "bonferroni")

# Compute Benjamini-Hochberg FDR values
t_test_results_df$BH_FDR_adjusted_p_value <- p.adjust(t_test_results_df$p_value, method = "BH") #using

#Final results of differential expression and p-value adjustments
View(t_test_results_df)

##### Filtering Bonferroni adjusted p values #####

# Filter rows based on Bonferroni-adjusted p-value with alpha = 0.01
filtered_df_bonferroni <- t_test_results_df[t_test_results_df$Bonferroni_adjusted_p_value < 0.01, ]
View(filtered_df_bonferroni)

# Create a new data frame without BH values
final_bonferroni_df <- data.frame(
  Probe = rownames(filtered_df_bonferroni),
  T_Stat = filtered_df_bonferroni$t_stat.t,
  p_value = filtered_df_bonferroni$p_value,
  Bonferroni_adjusted_p_value = filtered_df_bonferroni$Bonferroni_adjusted_p_value
)

# Display the new dataframe
View(final_bonferroni_df)

# Display the number of significant genes
cat("Number of genes passing Bonferroni correction with alpha = 0.01", ":", nrow(final_bonferroni_df),

```

```
## Number of genes passing Bonferroni correction with alpha = 0.01 : 3
```



```

#Write data frame into comma separated value file
write.csv(final_bonferroni_df, file = "DE_t_test_Bonferroni_alpha1%_Easwaran_Meena.csv", row.names = FALSE)

##### Filtering BH adjusted p values #####

# Filter rows based on Benjamini-Hochberg FDR-adjusted p-value with alpha = 0.01
filtered_df_BH <- t_test_results_df[t_test_results_df$BH_FDR_adjusted_p_value < 0.01, ]
View(filtered_df_BH)

# Create a new data frame without Bonferroni values
final_BH_df <- data.frame(
  Probe = rownames(filtered_df_BH),
  T_Stat = filtered_df_BH$t_stat.t,
  p_value = filtered_df_BH$p_value,
  BH_adjusted_p_value = filtered_df_BH$BH_FDR_adjusted_p_value
)

# Display the new dataframe
View(final_BH_df)

# Display the number of significant genes
cat("Number of genes passing Benjamini-Hochberg FDRcorrection with alpha = 0.01", ":", nrow(final_BH_df))

## Number of genes passing Benjamini-Hochberg FDRcorrection with alpha = 0.01 : 6

#Write data frame into comma separated value file
write.csv(final_BH_df, file = "DE_t_test_Benjamini_Hochberg_FDR_Easwaran_Meena.csv", row.names = FALSE)

```