

```

---
title: "TCGA LUSC using TCGA Biolinks"
output:
  pdf_document: default
  html_document: default
date: "2024-09-23"
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

## R Markdown


```{r cars}

# if (!requireNamespace("BiocManager", quietly = TRUE))
#install.packages("BiocManager")
#BiocManager::install("maftools")
#install.packages("pheatmap")
library(TCGAbiolinks)
library(tidyverse)
library(maftools)
library(pheatmap)
library(SummarizedExperiment)

#obtain project lists
gdcprojects <- getGDCprojects()
getProjectSummary('TCGA-LUSC')

# build a query
query_TCGA <- GDCquery(project = 'TCGA-LUSC', data.category = 'Transcriptome Profiling' )

output_query_TCGA <- getResults(query_TCGA)

# Build query to obtain gene expression data

query_TCGA <- GDCquery(project = 'TCGA-LUSC', data.category = 'Transcriptome Profiling',
access = 'open', experimental.strategy = 'RNA-Seq', workflow.type = 'STAR - Counts',
barcode = c('TCGA-43-7657-11A-01R-2125-07', 'TCGA-43-7657-01A-31R-2125-07', 'TCGA-60-2695-
01A-01R-0851-07') )

getResults(query_TCGA)

#Download data with GDCdownload

GDCdownload(query_TCGA)

#prepare data

tcga_lusc_data <- GDCprepare(query_TCGA, summarizedExperiment = TRUE)
lusc_matrix <- assay(tcga_lusc_data, 'unstranded' )
head(lusc_matrix)

# retrieve DNA methylation data
#Will only take top 3 cases

query_methyl <- GDCquery(project = 'TCGA-LUSC',
  data.category = 'DNA Methylation',

```

```

platform = 'Illumina Human Methylation 27',
access = 'open',
data.type = 'Methylation Beta Value',
barcode = c('TCGA-60-2710-01A-01D-0848-05', 'TCGA-22-1017-01A-01D-0689-05'))

output_query_methyl <- getResults(query_methyl)

GDCdownload(query_methyl)

#Plot probes showing differences in beta values
dna.meth <- GDCprepare(query_methyl, summarizedExperiment = TRUE)
head(assay(dna.meth))

idx <- dna.meth %>%
  assay %>%
  rowVars() %>%
  order(decreasing = TRUE) %>%
  head(10)

#plot
pheatmap(assay(dna.meth)[idx,])

```

```