

# Piecewise function to down-sample mutations in PCAWG spectra

Wu, Yang

2022-06-07

## 0. Load prerequisites

```
##      ggplot2
```

## 1. Definition of a piecewise function.

This function aims to lower the mutation counts in some hyper-mutated mutational spectra (e.g. spectra with Pol-Epsilon activity or UV exposure). On the other hand, it keeps the mutational spectra with normal or low mutation counts as they are.

```
source("down_sample.R")
print(down_sample_func) # Print the source code

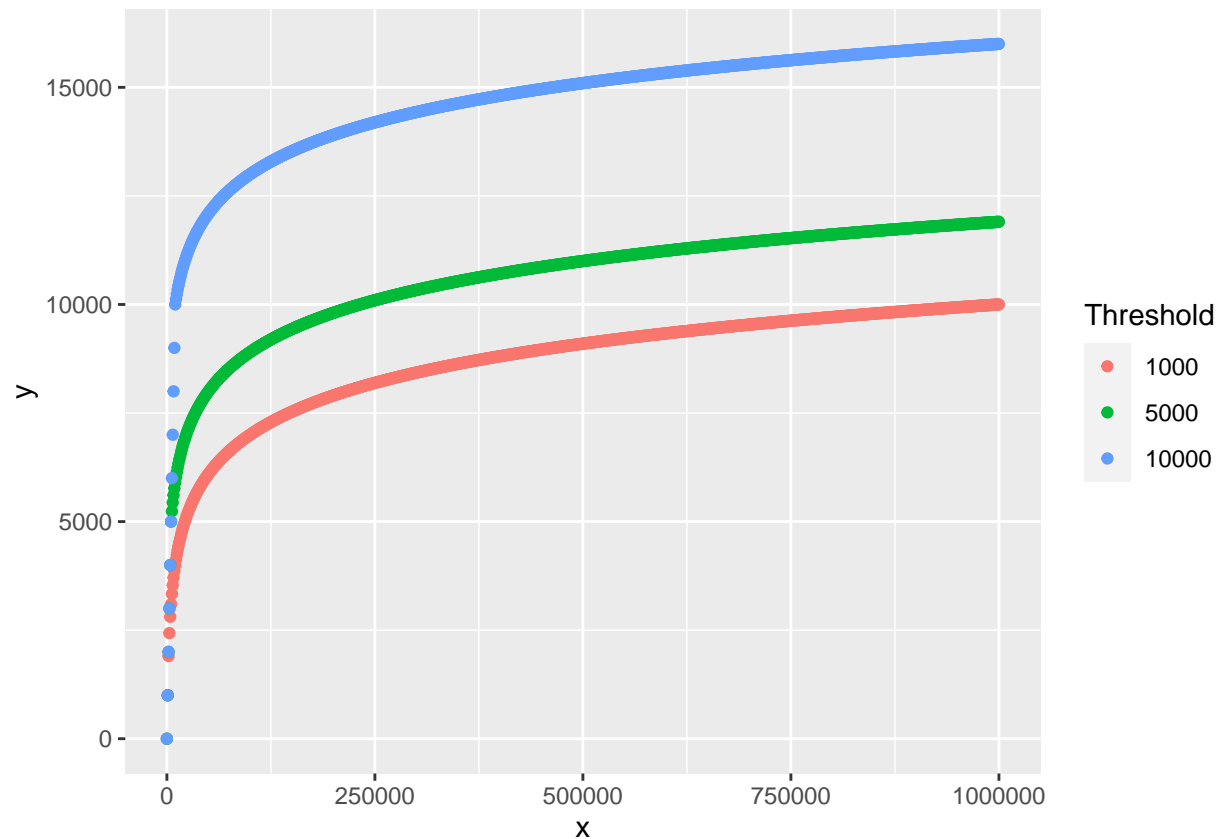
## function (x, thres = 1000L)
## {
##   if (x <= thres) {
##     return(x)
##   }
##   else {
##     return(ceiling(min(x, thres + 3000 * log10(x/thres))))
##   }
## }
```

## 2. Plotting the function

```
x_vec <- seq(0, 1e+06, 1000)
y_df <- data.frame()
thres_vals <- c(1000, 5000, 10000)
for (thres_val in thres_vals) {
  y_df <- rbind(
    y_df,
    data.frame(
      thres = thres_val,
      x = x_vec,
      y = sapply(x_vec, down_sample_func, thres = thres_val)))
}

ggObj <-
  ggplot2::ggplot(y_df, xlab = "Mutations",
                  ylab = "Transformed mutations") +
  geom_point(mapping = aes(x = x, y = y, color = as.factor(thres))) +
```

```
guides(color = guide_legend(title = "Threshold"))
ggobj
```



### 3. Print version

```
version
```

```
##
## platform      x86_64-w64-mingw32
## arch          x86_64
## os            mingw32
## crt           ucrt
## system        x86_64, mingw32
## status
## major         4
## minor         2.0
## year          2022
## month         04
## day           22
## svn rev       82229
## language      R
## version.string R version 4.2.0 (2022-04-22 ucrt)
## nickname      Vigorous Calisthenics
```