library(TxDb.Hsapiens.UCSC.hg38.knownGene)

library("GenomicFeatures")

##website: Index of /pub/release-108/gtf/homo\_sapiens (ensembl.org)

txdb <- TxDb.Hsapiens.UCSC.hg38.knownGene

txdb <- makeTxDbFromGFF("genome.gff3",format="gff3")

##get the exon position

exons\_gene <- exonsBy(txdb, by = "gene")

## remove the exon and calculate the exon length

exons\_gene\_len <- lapply(exons\_gene,function(x){sum(width(reduce(x)))})

##transfer into dataframe

# exons\_gene\_len\_df <- as.data.frame(t(as.data.frame(exons\_gene\_len)))

# exons\_gene\_len\_df$geneid <- gsub("X", "", rownames(exons\_gene\_len\_df))

gene\_length <- sapply(exons\_gene\_len,function(x){x})

id\_length <- as.data.frame(gene\_length)

# count to TPM

countToTpm <- function(counts, effLen){

rate <- log(counts) - log(effLen)

denom <- log(sum(exp(rate)))

exp(rate - denom + log(1e6))

}