Note Ryou Ready 3rd year

*#install.packages("tidyverse")*

**library**(tidyverse)

**library**(stringr)

**library**(htmlwidgets)

string\_Example <- "Molecular Biology, Cellular Biology, Cellular Oncology, Molecular Oncology"

*# Just show the first matches*

str\_view(string\_Example, "Biology")

str\_view\_all(string\_Example, "Biology")

str\_view\_all(string\_Example, c("Biology", "Oncology"))

mRNA <- "AUCCUAUCUCUAGACUAGAUCUAG"

str\_view\_all(mRNA, "CU")

mRNA <- "AUCCUAUCUCUAGACUAGAUCUAG"

str\_count(mRNA, "AG")

str\_count(mRNA, "CU")

str\_count(mRNA, c("AG", "CU"))

Location:

RNA <- "AUCCUAGUACUAGU"

str\_locate(RNA, "AGU") *#Location of 1st AGU*

str\_locate\_all(RNA , "AGU") *#Location of all AGUs*

str\_locate\_all(RNA , c("UAC","AGU")) *#Location of all AGUs*

Exercise 2.a and 2.b

Restriction Enzyme

Exp <- "Molecular Biology First Year Bachelor"

*#split by space*

str\_split(Exp, pattern = " ")

Example <- c("AB123CDE", "FG123HI")

Splitted <- str\_split(Example, pattern = "123")

Splitted

Splitted[[1]][1]

Splitted[[1]][2]

Splitted[[2]][1]

Splitted[[2]][2]

Exercise 2.c

Simulation

my\_dna <- c( "ACTGATCGATTACGTATAGTAGAATTCTATCATACATATATATCGATGCGTTCAT")

str\_count(my\_dna, "GAATTC")

str\_split(my\_dna, pattern = "GAATTC")

First\_Length <- str\_split(my\_dna, pattern = "GAATTC")[[1]][1]

First\_Length

str\_length(First\_Length)

Second\_Length <- str\_split(my\_dna, pattern = "GAATTC")[[1]][2]

Second\_Length

str\_length(Second\_Length)

*# Extract 1st element*

First\_Length <- str\_split(my\_dna, pattern = "GAATTC")[[1]][1]

First\_Length

*# Calculate 1st seq length*

str\_length(First\_Length) + 1

*# Extract 2nd string and 1st element*

Second\_Length <- str\_split(my\_dna, pattern = "GAATTC")[[1]][2]

Second\_Length

*# Calculate 2nd seq length*

str\_length(Second\_Length) + 5

Exercise 2.d, 3.e,

Splicing

Sample <- "123456789"

str\_sub(Sample, 4, 8)

str\_sub(Sample, 4)

DNA\_sample <- "AUCGCGUAGCGA"

Selected\_String <- str\_sub(DNA\_sample, 5, 8)

Selected\_String

Exercise 2.f:

Removing

DNA <- "ATCAAATCGATAGCTCGTA"

DNA

Remove1 <- str\_remove(DNA, "TCG")

Remove\_all <- str\_remove\_all(DNA, "TCG")