Mutation\_library\_analysis

Mu-En Chung

07/07/2022

[Introduction]  
This tool aligns a set of sample DNA sequences to a reference sequence,  
finds the mutations in each sample sequences, and then export a csv file with each row being each sample sequence and their amino acid substitutions.  
Author: Dr. Mu-En Chung Email: [chungmuen@gmail.com](mailto:chungmuen@gmail.com)

enter the path to the folder containing the sequencing results and the name of your WT sequence below:

seq\_file\_path = "./"   
ref\_seq\_name = "ClyF"

install\_and\_import\_packages:

if (!requireNamespace("BiocManager", quietly = TRUE)){  
 install.packages("BiocManager")  
}  
if (!requireNamespace("msa", quietly = TRUE)){  
 BiocManager::install("msa")  
}  
  
##import package msa  
library(msa)

## Loading required package: Biostrings

## 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, 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, sort, table, tapply,  
## union, unique, unsplit, which.max, which.min

## Loading required package: S4Vectors

## Loading required package: stats4

##   
## Attaching package: 'S4Vectors'

## The following objects are masked from 'package:base':  
##   
## expand.grid, I, unname

## Loading required package: IRanges

## Loading required package: XVector

## Loading required package: GenomeInfoDb

##   
## Attaching package: 'Biostrings'

## The following object is masked from 'package:base':  
##   
## strsplit

###run example provided by msa package  
mySequenceFile <- system.file("examples", "exampleAA.fasta", package="msa")  
mySequences <- readAAStringSet(mySequenceFile)  
print("example AA seq")

## [1] "example AA seq"

mySequences

## AAStringSet object of length 9:  
## width seq names   
## [1] 452 MSTAVLENPGLGRKLSDFGQETS...LKILADSINSEIGILCSALQKIK PH4H\_Homo\_sapiens  
## [2] 453 MAAVVLENGVLSRKLSDFGQETS...KILADSINSEVGILCNALQKIKS PH4H\_Rattus\_norve...  
## [3] 453 MAAVVLENGVLSRKLSDFGQETS...KILADSINSEVGILCHALQKIKS PH4H\_Mus\_musculus  
## [4] 297 MNDRADFVVPDITTRKNVGLSHD...DVAPDDLVLNAGDRQGWADTEDV PH4H\_Chromobacter...  
## [5] 262 MKTTQYVARQPDDNGFIHYPETE...MALVHEAMRLGLHAPLFPPKQAA PH4H\_Pseudomonas\_...  
## [6] 451 MSALVLESRALGRKLSDFGQETS...LKILADSISSEVEILCSALQKLK PH4H\_Bos\_taurus  
## [7] 313 MAIATPTSAAPTPAPAGFTGTLT...DVVDGDAVLNAGTREGWADTADI PH4H\_Ralstonia\_so...  
## [8] 294 MSGDGLSNGPPPGARPDWTIDQG...AVLTRGTQAYATAGGRLAGAAAG PH4H\_Caulobacter\_...  
## [9] 275 MSVAEYARDCAAQGLRGDYSVCR...QTADFEAIVARRKDQKALDPATV PH4H\_Rhizobium\_loti

myFirstAlignment <- msa(mySequences)

## use default substitution matrix

print("example AA alignment")

## [1] "example AA alignment"

myFirstAlignment

## CLUSTAL 2.1   
##   
## Call:  
## msa(mySequences)  
##   
## MsaAAMultipleAlignment with 9 rows and 456 columns  
## aln names  
## [1] MAAVVLENGVLSRKLSDFGQETSYIE...QLKILADSINSEVGILCNALQKIKS PH4H\_Rattus\_norve...  
## [2] MAAVVLENGVLSRKLSDFGQETSYIE...QLKILADSINSEVGILCHALQKIKS PH4H\_Mus\_musculus  
## [3] MSTAVLENPGLGRKLSDFGQETSYIE...QLKILADSINSEIGILCSALQKIK- PH4H\_Homo\_sapiens  
## [4] MSALVLESRALGRKLSDFGQETSYIE...QLKILADSISSEVEILCSALQKLK- PH4H\_Bos\_taurus  
## [5] --------------------------...LNAGDRQGWADTEDV---------- PH4H\_Chromobacter...  
## [6] --------------------------...LNAGTREGWADTADI---------- PH4H\_Ralstonia\_so...  
## [7] --------------------------...LTRGT-QAYATAGGRLAGAAAG--- PH4H\_Caulobacter\_...  
## [8] --------------------------...------------------------- PH4H\_Pseudomonas\_...  
## [9] --------------------------...------------------------- PH4H\_Rhizobium\_loti  
## Con --------------------------...??????????????IL??A???--- Consensus

##run example DNA alignment  
myDNAFile <- system.file("examples", "exampleDNA.fasta", package="msa")  
mySequencesDNA <- readDNAStringSet(myDNAFile)  
print("example DNA seq")

## [1] "example DNA seq"

mySequencesDNA

## DNAStringSet object of length 3:  
## width seq names   
## [1] 2680 CAGCTGGGGGTAAGGGGGGCGGA...AGTAATAAAACATTAGTAGTTTA PAH\_Homo\_sapiens  
## [2] 2152 AGTCCAGGACTGGAGTTGGGTAT...CATTATATAAAACAAAAAAAAAA PAH\_Mus\_musculus  
## [3] 2040 CAAGTTAAAACCATCAGCCCTCC...AAAAAAAAAAAAAAAAAAAAAAA PAH\_Rattus\_norveg...

exDNA\_Alignment <- msa(mySequencesDNA)

## use default substitution matrix

print("example DNA alignment")

## [1] "example DNA alignment"

exDNA\_Alignment

## CLUSTAL 2.1   
##   
## Call:  
## msa(mySequencesDNA)  
##   
## MsaDNAMultipleAlignment with 3 rows and 2716 columns  
## aln names  
## [1] --------------------------...------------------------- PAH\_Mus\_musculus  
## [2] --------------------------...------------------------- PAH\_Rattus\_norveg...  
## [3] CAGCTGGGGGTAAGGGGGGCGGATTA...TTAGTAATAAAACATTAGTAGTTTA PAH\_Homo\_sapiens  
## Con --------------------------...------------------------- Consensus

run your alignment

##run my alignment  
file\_path <- list.files(path = seq\_file\_path,pattern = 'fasta')  
seq\_result <- readDNAStringSet(file\_path)  
myDNA\_Alignment <- msa(seq\_result)

## use default substitution matrix

#The print() function provided by the msa package provides some ways for customizing the output, such as, showing the entire alignment split over multiple blocks of sub-sequences  
print("my alignment")

## [1] "my alignment"

print(myDNA\_Alignment, show="complete")

##   
## MsaDNAMultipleAlignment with 28 rows and 1296 columns  
## aln (1..53) names  
## [1] ------------------------------------------------AAAAT 31\_T7 Biotec\_CGN  
## [2] -------------------------------------------------GGAT 33\_T7 Biotec\_CGN  
## [3] ------------------------------------------------GAAGT 43\_T7 Biotec\_CGN  
## [4] -----------------------------------------------GGAAAA 21\_T7 Biotec\_CGN  
## [5] ------------------------------------------------GAAAT 51\_T7 Biotec\_CGN  
## [6] ------------------------------------------------AGGTT 14\_T7 Biotec\_CGN  
## [7] -AAGGGGAAATTCCTCTAGAAATAAATTTTTGTTAACTTTAAGAAGGAGATAT 55\_T7 Biotec\_CGN  
## [8] --GGGGGAAATTCCTCTAGAAATAAATTTTTGTTA-CTTTAAGAAGGAGATAT 56\_T7 Biotec\_CGN  
## [9] --GGGGGAAATTCCTCTAGAAATAA-TTTTTGTTA-CTTTAAGAAGGAGATAT 13\_T7 Biotec\_CGN   
## ... ...  
## [21] -CGGGGGAAATTTCTCTAGAAATAA-TTTTTGTTA-CTTTAAGAAGGAGATAT 23\_T7 Biotec\_CGN  
## [22] ---GGGGAAATTTCTCTAGAAATAA-TTTTTGTTA-CTTTAAGAAGGAGATAT 24\_T7 Biotec\_CGN  
## [23] --AGGGGAAATTCCTCTAGAAATAA-TTTTTGTTA-CTTTAAGAAGGAGATAT 25\_T7 Biotec\_CGN  
## [24] CAAGGGGAAATTCCTCTAGAAATAATTTTT-GTTTACTTTAAGAAGGAGATAT 32\_T7 Biotec\_CGN  
## [25] -------------------AAATAATTTTTTGTTTACTTTAAGAAGGAGATAT 34\_T7 Biotec\_CGN  
## [26] --AGGGGAAATTCCTCTAGAAATAA-TTTTTGTTA-CTTTAAGAAGGAGATAT 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] ----------------------------------------------GAAGT-T 45\_T7 Biotec\_CGN  
## Con --?GGGGAAATTCCTCTAGAAATAA-TTTTTGTTA-CTTTAAGAAGGAGATAT Consensus   
##   
## aln (54..106) names  
## [1] TCCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 31\_T7 Biotec\_CGN  
## [2] TACATGCATCACCATCACCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 33\_T7 Biotec\_CGN  
## [3] TACATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 43\_T7 Biotec\_CGN  
## [4] TACATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 21\_T7 Biotec\_CGN  
## [5] TCCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 51\_T7 Biotec\_CGN  
## [6] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 14\_T7 Biotec\_CGN  
## [7] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 55\_T7 Biotec\_CGN  
## [8] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 56\_T7 Biotec\_CGN  
## [9] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 23\_T7 Biotec\_CGN  
## [22] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 24\_T7 Biotec\_CGN  
## [23] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 25\_T7 Biotec\_CGN  
## [24] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 32\_T7 Biotec\_CGN  
## [25] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 34\_T7 Biotec\_CGN  
## [26] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 35\_T7 Biotec\_CGN  
## [27] ---ATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC ClyF  
## [28] ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC 45\_T7 Biotec\_CGN  
## Con ACCATGCATCACCATCATCACCACGCCCTTCCGAAAACGGGGAAACCTACAGC Consensus   
##   
## aln (107..159) names  
## [1] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 31\_T7 Biotec\_CGN  
## [2] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 33\_T7 Biotec\_CGN  
## [3] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 43\_T7 Biotec\_CGN  
## [4] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 21\_T7 Biotec\_CGN  
## [5] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 51\_T7 Biotec\_CGN  
## [6] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 14\_T7 Biotec\_CGN  
## [7] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 55\_T7 Biotec\_CGN  
## [8] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 56\_T7 Biotec\_CGN  
## [9] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 13\_T7 Biotec\_CGN   
## ... ...  
## [21] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 23\_T7 Biotec\_CGN  
## [22] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 24\_T7 Biotec\_CGN  
## [23] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 25\_T7 Biotec\_CGN  
## [24] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 32\_T7 Biotec\_CGN  
## [25] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 34\_T7 Biotec\_CGN  
## [26] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 35\_T7 Biotec\_CGN  
## [27] AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG ClyF  
## [28] AAGGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG 45\_T7 Biotec\_CGN  
## Con AAAGCAAGTGGTGGACTGGGCAATCAACTTGATCGGCTCCGGTGTTGACGTTG Consensus   
##   
## aln (160..212) names  
## [1] GCGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 31\_T7 Biotec\_CGN  
## [2] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 33\_T7 Biotec\_CGN  
## [3] ACGGGTATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 43\_T7 Biotec\_CGN  
## [4] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 21\_T7 Biotec\_CGN  
## [5] ACGGATATTACGGTCGTCAAAGTTGGGACTTGCCAAACTATATCTTTAACCGT 51\_T7 Biotec\_CGN  
## [6] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 14\_T7 Biotec\_CGN  
## [7] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 55\_T7 Biotec\_CGN  
## [8] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 56\_T7 Biotec\_CGN  
## [9] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 13\_T7 Biotec\_CGN   
## ... ...  
## [21] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 23\_T7 Biotec\_CGN  
## [22] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 24\_T7 Biotec\_CGN  
## [23] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 25\_T7 Biotec\_CGN  
## [24] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 32\_T7 Biotec\_CGN  
## [25] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 34\_T7 Biotec\_CGN  
## [26] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTCTAACCGT 35\_T7 Biotec\_CGN  
## [27] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT ClyF  
## [28] ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT 45\_T7 Biotec\_CGN  
## Con ACGGATATTACGGTCGTCAATGTTGGGACTTGCCAAACTATATCTTTAACCGT Consensus   
##   
## aln (213..265) names  
## [1] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 31\_T7 Biotec\_CGN  
## [2] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 33\_T7 Biotec\_CGN  
## [3] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 43\_T7 Biotec\_CGN  
## [4] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 21\_T7 Biotec\_CGN  
## [5] TACTGGAATTTCAAGACCCCAGGAAATACTCGCGACATGGCTTGGTATCGTTA 51\_T7 Biotec\_CGN  
## [6] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 14\_T7 Biotec\_CGN  
## [7] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 55\_T7 Biotec\_CGN  
## [8] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 56\_T7 Biotec\_CGN  
## [9] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 13\_T7 Biotec\_CGN   
## ... ...  
## [21] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 23\_T7 Biotec\_CGN  
## [22] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 24\_T7 Biotec\_CGN  
## [23] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 25\_T7 Biotec\_CGN  
## [24] TACTGGAATTTCGAGACCCCAGGAAATGCCCGCGACATGGCTTGGTATCGTTA 32\_T7 Biotec\_CGN  
## [25] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 34\_T7 Biotec\_CGN  
## [26] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 35\_T7 Biotec\_CGN  
## [27] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA ClyF  
## [28] TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA 45\_T7 Biotec\_CGN  
## Con TACTGGAATTTCAAGACCCCAGGAAATGCTCGCGACATGGCTTGGTATCGTTA Consensus   
##   
## aln (266..318) names  
## [1] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 31\_T7 Biotec\_CGN  
## [2] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 33\_T7 Biotec\_CGN  
## [3] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 43\_T7 Biotec\_CGN  
## [4] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 21\_T7 Biotec\_CGN  
## [5] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCGGCGATTTCGTACCAAAACCAG 51\_T7 Biotec\_CGN  
## [6] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 14\_T7 Biotec\_CGN  
## [7] CCCAGAGGGTTTCAAAGTTCTCCGCAACACCAGCGATTTCGTACCAAAACCAG 55\_T7 Biotec\_CGN  
## [8] CCCAGAGGGTTTCAAAGTTCTCCGCAACACCAGCGATTTCGTACCAAAACCAG 56\_T7 Biotec\_CGN  
## [9] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 13\_T7 Biotec\_CGN   
## ... ...  
## [21] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 23\_T7 Biotec\_CGN  
## [22] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 24\_T7 Biotec\_CGN  
## [23] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 25\_T7 Biotec\_CGN  
## [24] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTATCAAAACCAG 32\_T7 Biotec\_CGN  
## [25] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 34\_T7 Biotec\_CGN  
## [26] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 35\_T7 Biotec\_CGN  
## [27] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG ClyF  
## [28] CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG 45\_T7 Biotec\_CGN  
## Con CCCAGAGGGTTTCAAAGTTTTCCGCAACACCAGCGATTTCGTACCAAAACCAG Consensus   
##   
## aln (319..371) names  
## [1] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 31\_T7 Biotec\_CGN  
## [2] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 33\_T7 Biotec\_CGN  
## [3] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 43\_T7 Biotec\_CGN  
## [4] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 21\_T7 Biotec\_CGN  
## [5] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 51\_T7 Biotec\_CGN  
## [6] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 14\_T7 Biotec\_CGN  
## [7] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 55\_T7 Biotec\_CGN  
## [8] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 56\_T7 Biotec\_CGN  
## [9] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] GCGACATCGCGGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 23\_T7 Biotec\_CGN  
## [22] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 24\_T7 Biotec\_CGN  
## [23] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 25\_T7 Biotec\_CGN  
## [24] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 32\_T7 Biotec\_CGN  
## [25] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 34\_T7 Biotec\_CGN  
## [26] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC 35\_T7 Biotec\_CGN  
## [27] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC ClyF  
## [28] GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGAGGCAC 45\_T7 Biotec\_CGN  
## Con GCGACATCGCAGTGTGGACTGGGGGGAACTATAACTGGAATACATGGGGGCAC Consensus   
##   
## aln (372..424) names  
## [1] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 31\_T7 Biotec\_CGN  
## [2] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 33\_T7 Biotec\_CGN  
## [3] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 43\_T7 Biotec\_CGN  
## [4] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 21\_T7 Biotec\_CGN  
## [5] ACAGGA-TCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 51\_T7 Biotec\_CGN  
## [6] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 14\_T7 Biotec\_CGN  
## [7] ACAGGAATCGTCGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 55\_T7 Biotec\_CGN  
## [8] ACAGGAATCGTCGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 56\_T7 Biotec\_CGN  
## [9] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 13\_T7 Biotec\_CGN   
## ... ...  
## [21] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 23\_T7 Biotec\_CGN  
## [22] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 24\_T7 Biotec\_CGN  
## [23] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 25\_T7 Biotec\_CGN  
## [24] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGGCCA 32\_T7 Biotec\_CGN  
## [25] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 34\_T7 Biotec\_CGN  
## [26] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 35\_T7 Biotec\_CGN  
## [27] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA ClyF  
## [28] ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA 45\_T7 Biotec\_CGN  
## Con ACAGGAATCGTTGTTGGCCCAAGCACGAAGTCTTACTTCTATTCGGTTGACCA Consensus   
##   
## aln (425..477) names  
## [1] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 31\_T7 Biotec\_CGN  
## [2] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 33\_T7 Biotec\_CGN  
## [3] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 43\_T7 Biotec\_CGN  
## [4] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 21\_T7 Biotec\_CGN  
## [5] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 51\_T7 Biotec\_CGN  
## [6] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 14\_T7 Biotec\_CGN  
## [7] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 55\_T7 Biotec\_CGN  
## [8] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 56\_T7 Biotec\_CGN  
## [9] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 23\_T7 Biotec\_CGN  
## [22] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 24\_T7 Biotec\_CGN  
## [23] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 25\_T7 Biotec\_CGN  
## [24] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 32\_T7 Biotec\_CGN  
## [25] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 34\_T7 Biotec\_CGN  
## [26] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 35\_T7 Biotec\_CGN  
## [27] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC ClyF  
## [28] AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC 45\_T7 Biotec\_CGN  
## Con AAATTGGAACAACAGTAACAGTTATGTCGGCAGTCCTGCCGCCAAAATCAAGC Consensus   
##   
## aln (478..530) names  
## [1] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 31\_T7 Biotec\_CGN  
## [2] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 33\_T7 Biotec\_CGN  
## [3] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 43\_T7 Biotec\_CGN  
## [4] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 21\_T7 Biotec\_CGN  
## [5] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 51\_T7 Biotec\_CGN  
## [6] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 14\_T7 Biotec\_CGN  
## [7] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 55\_T7 Biotec\_CGN  
## [8] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 56\_T7 Biotec\_CGN  
## [9] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 13\_T7 Biotec\_CGN   
## ... ...  
## [21] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 23\_T7 Biotec\_CGN  
## [22] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 24\_T7 Biotec\_CGN  
## [23] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 25\_T7 Biotec\_CGN  
## [24] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 32\_T7 Biotec\_CGN  
## [25] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 34\_T7 Biotec\_CGN  
## [26] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 35\_T7 Biotec\_CGN  
## [27] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG ClyF  
## [28] ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG 45\_T7 Biotec\_CGN  
## Con ACTCCTACTTTGGAGTGACCCATTTTGTGCGTCCCGCTTACAAAGCCGAACCG Consensus   
##   
## aln (531..583) names  
## [1] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 31\_T7 Biotec\_CGN  
## [2] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 33\_T7 Biotec\_CGN  
## [3] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 43\_T7 Biotec\_CGN  
## [4] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 21\_T7 Biotec\_CGN  
## [5] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 51\_T7 Biotec\_CGN  
## [6] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 14\_T7 Biotec\_CGN  
## [7] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 55\_T7 Biotec\_CGN  
## [8] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 56\_T7 Biotec\_CGN  
## [9] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 23\_T7 Biotec\_CGN  
## [22] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 24\_T7 Biotec\_CGN  
## [23] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 25\_T7 Biotec\_CGN  
## [24] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 32\_T7 Biotec\_CGN  
## [25] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 34\_T7 Biotec\_CGN  
## [26] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 35\_T7 Biotec\_CGN  
## [27] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC ClyF  
## [28] AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC 45\_T7 Biotec\_CGN  
## Con AAGCCTACTCCGCCATCTCGCAGTCGTTCATACCGCGAAACAGGAACGATGAC Consensus   
##   
## aln (584..636) names  
## [1] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 31\_T7 Biotec\_CGN  
## [2] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 33\_T7 Biotec\_CGN  
## [3] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 43\_T7 Biotec\_CGN  
## [4] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 21\_T7 Biotec\_CGN  
## [5] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 51\_T7 Biotec\_CGN  
## [6] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 14\_T7 Biotec\_CGN  
## [7] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 55\_T7 Biotec\_CGN  
## [8] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 56\_T7 Biotec\_CGN  
## [9] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 13\_T7 Biotec\_CGN   
## ... ...  
## [21] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 23\_T7 Biotec\_CGN  
## [22] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 24\_T7 Biotec\_CGN  
## [23] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 25\_T7 Biotec\_CGN  
## [24] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 32\_T7 Biotec\_CGN  
## [25] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 34\_T7 Biotec\_CGN  
## [26] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 35\_T7 Biotec\_CGN  
## [27] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA ClyF  
## [28] GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA 45\_T7 Biotec\_CGN  
## Con GGTCACGGTGGATGCGCTGAACGTACGTCGCGCTCCGAACACCTCGGGTGAAA Consensus   
##   
## aln (637..689) names  
## [1] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCGTCATC 31\_T7 Biotec\_CGN  
## [2] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 33\_T7 Biotec\_CGN  
## [3] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 43\_T7 Biotec\_CGN  
## [4] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 21\_T7 Biotec\_CGN  
## [5] TTGTGGCTGTGTACAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 51\_T7 Biotec\_CGN  
## [6] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 14\_T7 Biotec\_CGN  
## [7] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 55\_T7 Biotec\_CGN  
## [8] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 56\_T7 Biotec\_CGN  
## [9] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 23\_T7 Biotec\_CGN  
## [22] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 24\_T7 Biotec\_CGN  
## [23] TTGTGGCTGTGTTTAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 25\_T7 Biotec\_CGN  
## [24] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 32\_T7 Biotec\_CGN  
## [25] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 34\_T7 Biotec\_CGN  
## [26] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 35\_T7 Biotec\_CGN  
## [27] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC ClyF  
## [28] TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC 45\_T7 Biotec\_CGN  
## Con TTGTGGCTGTGTATAAACGCGGTGAGAGCTTCGACTATGATACGGTCATCATC Consensus   
##   
## aln (690..742) names  
## [1] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 31\_T7 Biotec\_CGN  
## [2] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 33\_T7 Biotec\_CGN  
## [3] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 43\_T7 Biotec\_CGN  
## [4] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 21\_T7 Biotec\_CGN  
## [5] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGGAGCGCAA 51\_T7 Biotec\_CGN  
## [6] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 14\_T7 Biotec\_CGN  
## [7] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 55\_T7 Biotec\_CGN  
## [8] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 56\_T7 Biotec\_CGN  
## [9] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 13\_T7 Biotec\_CGN   
## ... ...  
## [21] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 23\_T7 Biotec\_CGN  
## [22] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 24\_T7 Biotec\_CGN  
## [23] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 25\_T7 Biotec\_CGN  
## [24] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 32\_T7 Biotec\_CGN  
## [25] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 34\_T7 Biotec\_CGN  
## [26] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 35\_T7 Biotec\_CGN  
## [27] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA ClyF  
## [28] GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA 45\_T7 Biotec\_CGN  
## Con GACGTGAATGGTTACGTCTGGGTGTCTTACATTGGCGGGTCCGGGAAGCGCAA Consensus   
##   
## aln (743..795) names  
## [1] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 31\_T7 Biotec\_CGN  
## [2] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 33\_T7 Biotec\_CGN  
## [3] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 43\_T7 Biotec\_CGN  
## [4] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 21\_T7 Biotec\_CGN  
## [5] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 51\_T7 Biotec\_CGN  
## [6] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 14\_T7 Biotec\_CGN  
## [7] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 55\_T7 Biotec\_CGN  
## [8] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 56\_T7 Biotec\_CGN  
## [9] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 13\_T7 Biotec\_CGN   
## ... ...  
## [21] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 23\_T7 Biotec\_CGN  
## [22] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 24\_T7 Biotec\_CGN  
## [23] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 25\_T7 Biotec\_CGN  
## [24] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 32\_T7 Biotec\_CGN  
## [25] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 34\_T7 Biotec\_CGN  
## [26] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 35\_T7 Biotec\_CGN  
## [27] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG ClyF  
## [28] CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG 45\_T7 Biotec\_CGN  
## Con CTACGTTGCAACAGGGGCAACGAAGGACGGCAAGCGTTTTGGAAATGCATGGG Consensus   
##   
## aln (796..848) names  
## [1] G-TACTTTTAAGTAATAAGAA-------------------------------- 31\_T7 Biotec\_CGN  
## [2] G-TACTTTTAAGTAATAAGAA-------------------------------- 33\_T7 Biotec\_CGN  
## [3] G-TACTTTTAAGTAATAAGAA-------------------------------- 43\_T7 Biotec\_CGN  
## [4] G-TACTTTTAAGTAATAAGAA-------------------------------- 21\_T7 Biotec\_CGN  
## [5] G-TACTTTTAAGTAATAAGAA-------------------------------- 51\_T7 Biotec\_CGN  
## [6] G-TACTTTTAAGTAATAAGAA-------------------------------- 14\_T7 Biotec\_CGN  
## [7] G-TACTTTTAAGTAATAAGAA-------------------------------- 55\_T7 Biotec\_CGN  
## [8] G-TACTTTTAAGTAATAAGAA-------------------------------- 56\_T7 Biotec\_CGN  
## [9] G-TACTTTTAAGTAATAAGAA-------------------------------- 13\_T7 Biotec\_CGN   
## ... ...  
## [21] G-TACTTTTAAGTAATAAGAA-------------------------------- 23\_T7 Biotec\_CGN  
## [22] G-TACTTTTAAGTAATAAGAA-------------------------------- 24\_T7 Biotec\_CGN  
## [23] G-TACTTTTAAGTAATAAGAA-------------------------------- 25\_T7 Biotec\_CGN  
## [24] G-TACTTTTAAGTAATAAGAA-------------------------------- 32\_T7 Biotec\_CGN  
## [25] G-TACTTTTAAGTAATAAGAA-------------------------------- 34\_T7 Biotec\_CGN  
## [26] G-TACTTTTAAGTAATAAGAA-------------------------------- 35\_T7 Biotec\_CGN  
## [27] G-TACTTTTAAGTAA-------------------------------------- ClyF  
## [28] G-TACTTTTAAGTAATAAGAATTCGAGCTCCGTCGACAAGCTTGCGGGCCGCA 45\_T7 Biotec\_CGN  
## Con G-TACTTTTAAGTAATAAGAA-------------------------------- Consensus   
##   
## aln (849..901) names  
## [1] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 31\_T7 Biotec\_CGN  
## [2] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 33\_T7 Biotec\_CGN  
## [3] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 43\_T7 Biotec\_CGN  
## [4] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 21\_T7 Biotec\_CGN  
## [5] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 51\_T7 Biotec\_CGN  
## [6] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 14\_T7 Biotec\_CGN  
## [7] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 55\_T7 Biotec\_CGN  
## [8] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 56\_T7 Biotec\_CGN  
## [9] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 23\_T7 Biotec\_CGN  
## [22] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 24\_T7 Biotec\_CGN  
## [23] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 25\_T7 Biotec\_CGN  
## [24] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 32\_T7 Biotec\_CGN  
## [25] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 34\_T7 Biotec\_CGN  
## [26] TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] CTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC 45\_T7 Biotec\_CGN  
## Con TTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACC-ACCAC Consensus   
##   
## aln (902..954) names  
## [1] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 31\_T7 Biotec\_CGN  
## [2] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 33\_T7 Biotec\_CGN  
## [3] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 43\_T7 Biotec\_CGN  
## [4] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 21\_T7 Biotec\_CGN  
## [5] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 51\_T7 Biotec\_CGN  
## [6] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 14\_T7 Biotec\_CGN  
## [7] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 55\_T7 Biotec\_CGN  
## [8] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 56\_T7 Biotec\_CGN  
## [9] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 13\_T7 Biotec\_CGN   
## ... ...  
## [21] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 23\_T7 Biotec\_CGN  
## [22] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 24\_T7 Biotec\_CGN  
## [23] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 25\_T7 Biotec\_CGN  
## [24] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 32\_T7 Biotec\_CGN  
## [25] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 34\_T7 Biotec\_CGN  
## [26] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT 45\_T7 Biotec\_CGN  
## Con CACTGA-GATCCGG-CTGCTAACAAAGCCC---GAAAGGAAGCTGAGTTGGCT Consensus   
##   
## aln (955..1007) names  
## [1] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 31\_T7 Biotec\_CGN  
## [2] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 33\_T7 Biotec\_CGN  
## [3] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 43\_T7 Biotec\_CGN  
## [4] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 21\_T7 Biotec\_CGN  
## [5] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 51\_T7 Biotec\_CGN  
## [6] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 14\_T7 Biotec\_CGN  
## [7] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 55\_T7 Biotec\_CGN  
## [8] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 56\_T7 Biotec\_CGN  
## [9] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 13\_T7 Biotec\_CGN   
## ... ...  
## [21] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 23\_T7 Biotec\_CGN  
## [22] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 24\_T7 Biotec\_CGN  
## [23] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 25\_T7 Biotec\_CGN  
## [24] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 32\_T7 Biotec\_CGN  
## [25] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 34\_T7 Biotec\_CGN  
## [26] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA 45\_T7 Biotec\_CGN  
## Con GCTGCCACC-GCTGAGCAATAACT--AGCATAACCCCTT--GGGGCCTCTAAA Consensus   
##   
## aln (1008..1060) names  
## [1] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 31\_T7 Biotec\_CGN  
## [2] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 33\_T7 Biotec\_CGN  
## [3] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 43\_T7 Biotec\_CGN  
## [4] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 21\_T7 Biotec\_CGN  
## [5] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 51\_T7 Biotec\_CGN  
## [6] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 14\_T7 Biotec\_CGN  
## [7] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 55\_T7 Biotec\_CGN  
## [8] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 56\_T7 Biotec\_CGN  
## [9] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 13\_T7 Biotec\_CGN   
## ... ...  
## [21] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 23\_T7 Biotec\_CGN  
## [22] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 24\_T7 Biotec\_CGN  
## [23] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 25\_T7 Biotec\_CGN  
## [24] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 32\_T7 Biotec\_CGN  
## [25] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 34\_T7 Biotec\_CGN  
## [26] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT 45\_T7 Biotec\_CGN  
## Con C--GGGTCTTGAGGGG---TTTTTTGCTGAAA--GGAGGAACTATATCCGGAT Consensus   
##   
## aln (1061..1113) names  
## [1] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 31\_T7 Biotec\_CGN  
## [2] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 33\_T7 Biotec\_CGN  
## [3] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 43\_T7 Biotec\_CGN  
## [4] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 21\_T7 Biotec\_CGN  
## [5] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 51\_T7 Biotec\_CGN  
## [6] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 14\_T7 Biotec\_CGN  
## [7] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 55\_T7 Biotec\_CGN  
## [8] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 56\_T7 Biotec\_CGN  
## [9] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 13\_T7 Biotec\_CGN   
## ... ...  
## [21] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 23\_T7 Biotec\_CGN  
## [22] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 24\_T7 Biotec\_CGN  
## [23] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 25\_T7 Biotec\_CGN  
## [24] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 32\_T7 Biotec\_CGN  
## [25] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 34\_T7 Biotec\_CGN  
## [26] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG 45\_T7 Biotec\_CGN  
## Con TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG Consensus   
##   
## aln (1114..1166) names  
## [1] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 31\_T7 Biotec\_CGN  
## [2] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 33\_T7 Biotec\_CGN  
## [3] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 43\_T7 Biotec\_CGN  
## [4] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 21\_T7 Biotec\_CGN  
## [5] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 51\_T7 Biotec\_CGN  
## [6] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 14\_T7 Biotec\_CGN  
## [7] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 55\_T7 Biotec\_CGN  
## [8] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 56\_T7 Biotec\_CGN  
## [9] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 13\_T7 Biotec\_CGN   
## ... ...  
## [21] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 23\_T7 Biotec\_CGN  
## [22] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 24\_T7 Biotec\_CGN  
## [23] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 25\_T7 Biotec\_CGN  
## [24] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 32\_T7 Biotec\_CGN  
## [25] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 34\_T7 Biotec\_CGN  
## [26] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC 45\_T7 Biotec\_CGN  
## Con TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTC Consensus   
##   
## aln (1167..1219) names  
## [1] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 31\_T7 Biotec\_CGN  
## [2] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 33\_T7 Biotec\_CGN  
## [3] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 43\_T7 Biotec\_CGN  
## [4] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 21\_T7 Biotec\_CGN  
## [5] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 51\_T7 Biotec\_CGN  
## [6] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 14\_T7 Biotec\_CGN  
## [7] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGG-CTTTCCCCGTCAAGCTC 55\_T7 Biotec\_CGN  
## [8] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGGTTTTCCCCGTCAAGCTC 56\_T7 Biotec\_CGN  
## [9] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCT---------------- 13\_T7 Biotec\_CGN   
## ... ...  
## [21] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGGTTTTCCCCGTCAAGCTC 23\_T7 Biotec\_CGN  
## [22] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGGTTTTCCCCGTCAAGCTC 24\_T7 Biotec\_CGN  
## [23] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGGTTTTCCCCGTCAAGCTC 25\_T7 Biotec\_CGN  
## [24] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGGTTTTCCCCGTCAAGCTC 32\_T7 Biotec\_CGN  
## [25] GCTTTCTTCCCTTCCTTTCTCGC------------------------------ 34\_T7 Biotec\_CGN  
## [26] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTC- 45\_T7 Biotec\_CGN  
## Con GCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGG?TTT?CCC????A???C? Consensus   
##   
## aln (1220..1272) names  
## [1] AAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACC 31\_T7 Biotec\_CGN  
## [2] AAATCGGGGGCTCCCTTTAGGGTTCCAATTTAGTGCTTTACGGCACCTCGACC 33\_T7 Biotec\_CGN  
## [3] AAATCGGGGGCTCCCTTTAGGGTTCCAATTTATTGCTTTACGGCACCTC---- 43\_T7 Biotec\_CGN  
## [4] AAATCGGGGGCTCCCTTTAGGGTTCCAATT-AGTGCTTTACGGCCC------- 21\_T7 Biotec\_CGN  
## [5] AAATCGGGGGCTCCCTTTAGGGTTCCAATTTAATGCTTTACGGCACCTC---- 51\_T7 Biotec\_CGN  
## [6] AAATCGGGGGCTCCCTTTAGGGTTCCAATTTA--------------------- 14\_T7 Biotec\_CGN  
## [7] TAAATCGGGGGCTCC-------------------------------------- 55\_T7 Biotec\_CGN  
## [8] TAAATCGGGGGCTCCCTTTAGGGTTCCCAATTTA------------------- 56\_T7 Biotec\_CGN  
## [9] ----------------------------------------------------- 13\_T7 Biotec\_CGN   
## ... ...  
## [21] TAAATCGGGGGCTCCCTTTAGGGTTCCA------------------------- 23\_T7 Biotec\_CGN  
## [22] TAAATCGGGGGCTCCCTTTAGGGTTCC-------------------------- 24\_T7 Biotec\_CGN  
## [23] TAAATCGGGGGCTCCCTTTAGGGTTCCAATTTA-------------------- 25\_T7 Biotec\_CGN  
## [24] TAAATCGGGGGCTCCCTTTAGGGTTCCGAATTTAGG----------------- 32\_T7 Biotec\_CGN  
## [25] ----------------------------------------------------- 34\_T7 Biotec\_CGN  
## [26] AAATCGGGGGCTCCCTTTAGGGTTCCAA------------------------- 35\_T7 Biotec\_CGN  
## [27] ----------------------------------------------------- ClyF  
## [28] ----------------------------------------------------- 45\_T7 Biotec\_CGN  
## Con ?AA???GGGG??-??-------------------------------------- Consensus   
##   
## aln (1273..1296) names  
## [1] CCAAAAACTTGATTAGGGTGAGGG 31\_T7 Biotec\_CGN  
## [2] CCAAAAACTTGATTAGGGT----- 33\_T7 Biotec\_CGN  
## [3] ------------------------ 43\_T7 Biotec\_CGN  
## [4] ------------------------ 21\_T7 Biotec\_CGN  
## [5] ------------------------ 51\_T7 Biotec\_CGN  
## [6] ------------------------ 14\_T7 Biotec\_CGN  
## [7] ------------------------ 55\_T7 Biotec\_CGN  
## [8] ------------------------ 56\_T7 Biotec\_CGN  
## [9] ------------------------ 13\_T7 Biotec\_CGN   
## ... ...  
## [21] ------------------------ 23\_T7 Biotec\_CGN  
## [22] ------------------------ 24\_T7 Biotec\_CGN  
## [23] ------------------------ 25\_T7 Biotec\_CGN  
## [24] ------------------------ 32\_T7 Biotec\_CGN  
## [25] ------------------------ 34\_T7 Biotec\_CGN  
## [26] ------------------------ 35\_T7 Biotec\_CGN  
## [27] ------------------------ ClyF  
## [28] ------------------------ 45\_T7 Biotec\_CGN  
## Con ------------------------ Consensus

from alignment, calculate the mutation profile in DNA level and in Amino acid level

#convert alignment to a list of vectors and matrix   
convert <- msaConvert(myDNA\_Alignment,type= "bio3d::fasta")  
  
alignment\_mat <- convert$ali  
alignment\_df <- as.data.frame(alignment\_mat)  
  
alignment\_df\_dREF <- alignment\_df[-grep(ref\_seq\_name,row.names(alignment\_df)),]  
alignment\_df\_dREF <- alignment\_df\_dREF[order(row.names(alignment\_df\_dREF)),]  
  
alignment\_df\_reorder <- as.data.frame(rbind(alignment\_df[grep(ref\_seq\_name,row.names(alignment\_df)),],alignment\_df\_dREF))  
  
#remove columns extruding the reference sequence  
keep\_columns <- as.data.frame(which(alignment\_df\_reorder[1,]!= "-"))  
trim\_before <- keep\_columns[1,1]-1  
trim\_after <- keep\_columns[length(keep\_columns[,1]),1]+1  
trim\_alignment <- alignment\_df\_reorder[,-c(1:trim\_before,trim\_after:ncol(alignment\_df\_reorder))]  
# trim\_alignment[1,]  
  
#identify the position of "-" in the reference   
reference\_row <- trim\_alignment[1,]  
print("afer alignment, the positions of - in the reference sequence are:")

## [1] "afer alignment, the positions of - in the reference sequence are:"

which(reference\_row == "-")

## [1] 741

#identify the mutations and create a matrix of mutations (0=mutation, 1=no mutation)  
mutation\_count <- trim\_alignment  
  
for (i in 1:nrow(trim\_alignment))   
{  
 for (j in 1:ncol(trim\_alignment))   
 {  
 if (trim\_alignment[i, j] == trim\_alignment[1, j])   
 {  
 mutation\_count[i, j] <- 1  
 }   
 else   
 {  
 mutation\_count[i, j] <- 0  
 }  
 }  
}  
  
################################################  
  
#classfy the mutations and create a matrix of mutations  
mutation\_count\_TvTs <- trim\_alignment  
  
for (i in 1:nrow(trim\_alignment))   
{  
 for (j in 1:ncol(trim\_alignment))   
 {  
 if (trim\_alignment[i,j] == trim\_alignment[1,j])   
 {  
 mutation\_count\_TvTs[i,j] <- "-"  
 }   
 else if (trim\_alignment[i,j] == "A")  
 {  
 if (trim\_alignment[1,j] == "T")  
 {  
 mutation\_count\_TvTs[i, j] <- "TA"  
 }  
 else if (trim\_alignment[1,j] == "C")  
 {  
 mutation\_count\_TvTs[i, j] <- "CA"  
 }  
 else if (trim\_alignment[1,j] == "G")  
 {  
 mutation\_count\_TvTs[i, j] <- "GA"  
 }  
 else if (trim\_alignment[1,j] == "-")  
 {  
 mutation\_count\_TvTs[i, j] <- "Ins"  
 }  
 }  
 else if (trim\_alignment[i,j] == "T")  
 {  
 if (trim\_alignment[1,j] == "A")  
 {  
 mutation\_count\_TvTs[i, j] <- "AT"  
 }  
 else if (trim\_alignment[1,j] == "C")  
 {  
 mutation\_count\_TvTs[i, j] <- "CT"  
 }  
 else if (trim\_alignment[1,j] == "G")  
 {  
 mutation\_count\_TvTs[i, j] <- "GT"  
 }  
 else if (trim\_alignment[1,j] == "-")  
 {  
 mutation\_count\_TvTs[i, j] <- "Ins"  
 }  
 }  
 else if (trim\_alignment[i,j] == "C")  
 {  
 if (trim\_alignment[1,j] == "A")  
 {  
 mutation\_count\_TvTs[i, j] <- "AC"  
 }  
 else if (trim\_alignment[1,j] == "T")  
 {  
 mutation\_count\_TvTs[i, j] <- "TC"  
 }  
 else if (trim\_alignment[1,j] == "G")  
 {  
 mutation\_count\_TvTs[i, j] <- "GC"  
 }  
 else if (trim\_alignment[1,j] == "-")  
 {  
 mutation\_count\_TvTs[i, j] <- "Ins"  
 }  
 }  
 else if (trim\_alignment[i,j] == "G")  
 {  
 if (trim\_alignment[1,j] == "A")  
 {  
 mutation\_count\_TvTs[i, j] <- "AG"  
 }  
 else if (trim\_alignment[1,j] == "T")  
 {  
 mutation\_count\_TvTs[i, j] <- "TG"  
 }  
 else if (trim\_alignment[1,j] == "C")  
 {  
 mutation\_count\_TvTs[i, j] <- "CG"  
 }  
 else if (trim\_alignment[1,j] == "-")  
 {  
 mutation\_count\_TvTs[i, j] <- "Ins"  
 }  
 }  
 else if (trim\_alignment[i,j] == "-")  
 {  
 mutation\_count\_TvTs[i, j] <- "Del"  
 }   
 else  
 {  
 mutation\_count\_TvTs[i, j] <- "?"  
 }   
 }  
}  
  
#check if all the elements are assigned as "-", "S", "V", "Del", "Ins", "?"  
print('the number of "-", "S", "V", "Del", "Ins", "?":')

## [1] "the number of \"-\", \"S\", \"V\", \"Del\", \"Ins\", \"?\":"

for(i in 1:nrow(mutation\_count\_TvTs))  
{  
 a <- table(as.character(mutation\_count\_TvTs[i,]))  
 print(row.names(mutation\_count\_TvTs)[i])  
 print(a)  
}

## [1] "ClyF"  
##   
## -   
## 754   
## [1] "11\_T7 Biotec\_CGN"  
##   
## - AG   
## 753 1   
## [1] "12\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "13\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "14\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "15\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "21\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "22\_T7 Biotec\_CGN"  
##   
## - Ins   
## 753 1   
## [1] "23\_T7 Biotec\_CGN"  
##   
## - AG   
## 753 1   
## [1] "24\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "25\_T7 Biotec\_CGN"  
##   
## - AT   
## 753 1   
## [1] "31\_T7 Biotec\_CGN"  
##   
## - AG   
## 752 2   
## [1] "32\_T7 Biotec\_CGN"  
##   
## - AG CT TC   
## 750 2 1 1   
## [1] "33\_T7 Biotec\_CGN"  
##   
## - TC   
## 753 1   
## [1] "34\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "35\_T7 Biotec\_CGN"  
##   
## - TC   
## 753 1   
## [1] "41\_T7 Biotec\_CGN"  
##   
## - CT   
## 753 1   
## [1] "42\_T7 Biotec\_CGN"  
##   
## - CT GA TC   
## 751 1 1 1   
## [1] "43\_T7 Biotec\_CGN"  
##   
## - AG   
## 753 1   
## [1] "44\_T7 Biotec\_CGN"  
##   
## - TA TC   
## 751 1 2   
## [1] "45\_T7 Biotec\_CGN"  
##   
## - AG GA   
## 752 1 1   
## [1] "46\_T7 Biotec\_CGN"  
##   
## - AG GA TC   
## 751 1 1 1   
## [1] "51\_T7 Biotec\_CGN"  
##   
## - AG Del GA TA TC   
## 748 2 1 1 1 1   
## [1] "52\_T7 Biotec\_CGN"  
##   
## -   
## 754   
## [1] "53\_T7 Biotec\_CGN"  
##   
## - AG CT TC   
## 750 1 2 1   
## [1] "54\_T7 Biotec\_CGN"  
##   
## - AG TC   
## 751 1 2   
## [1] "55\_T7 Biotec\_CGN"  
##   
## - TC   
## 752 2   
## [1] "56\_T7 Biotec\_CGN"  
##   
## - TC   
## 752 2

#show row names  
rownames <- as.data.frame(row.names(mutation\_count))  
  
#count the mutations in each sequence  
#output:   
#id - S V Ins Del ?  
mutation\_count\_row <- t(as.matrix(table(as.character(mutation\_count\_TvTs[1,]))))  
addition <- matrix(c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0), nrow = 1, ncol = 15)  
mutation\_count\_row <- as.data.frame(cbind(mutation\_count\_row,addition))  
colnames(mutation\_count\_row) <- c("-","AG", "GA", "AT", "TA", "AC", "CA", "GC", "CG", "GT", "TG", "CT", "TC", "Del", "Ins", "?")  
  
for(i in 2:nrow(mutation\_count\_TvTs))  
{  
 mutation\_count\_row2 <- as.data.frame(t(as.matrix(table(as.character(mutation\_count\_TvTs[i,])))))  
 if(!('AG' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'AG'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('GA' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'GA'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('AT' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'AT'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('TA' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'TA'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('AC' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'AC'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('CA' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'CA'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('GC' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'GC'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('CG' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'CG'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('GT' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'GT'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('TG' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'TG'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('CT' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'CT'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('TC' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'TC'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('Del' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'Del'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('Ins' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- 'Ins'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 if(!('?' %in% names(mutation\_count\_row2)))  
 {  
 addition <- as.data.frame(matrix(0, nrow = 1, ncol = 1))  
 colnames(addition) <- '?'  
 mutation\_count\_row2 <- as.data.frame(cbind(mutation\_count\_row2,addition))  
 }  
 mutation\_count\_row2  
   
 mutation\_count\_row <- rbind(mutation\_count\_row,mutation\_count\_row2)  
}  
row.names(mutation\_count\_row) <- row.names(mutation\_count\_TvTs)  
  
################################################  
#identify the row with the most mutations  
n\_mut <- sum(mutation\_count[1,]==0)  
for(i in 2:nrow(mutation\_count))  
{  
 n\_mut2 <- sum(mutation\_count[i,]==0)  
 if(n\_mut2 >= n\_mut)  
 {  
 n\_mut <- n\_mut2  
 }  
}  
###############################################  
#identify mutation locations  
mutation\_locations <- rep(NA,n\_mut)  
  
for(i in 2:nrow(mutation\_count))  
{  
 mutation\_locations2 <- rep(NA,n\_mut)  
 mutations <- t(as.data.frame(which(mutation\_count[i,]==0)))  
 for (j in 1:n\_mut)  
 {  
 if (j <= ncol(mutations))  
 {  
 mutation\_locations2 [j] <- mutations[1,j]  
 }  
 }  
 mutation\_locations <- rbind(mutation\_locations, mutation\_locations2)  
}  
row.names(mutation\_locations) <- row.names(mutation\_count)  
  
###########################################################################  
#identify the amino acid mutations  
if(sum(trim\_alignment[1,]=="-")!= 0)  
{  
 trim\_dash <- trim\_alignment[,-c(which(trim\_alignment[1,]=="-"))]   
}else  
{  
 trim\_dash <- trim\_alignment  
}  
  
modulo\_mut\_loc\_df <- mutation\_locations   
AAmutations\_df <- mutation\_locations  
for(i in 1:nrow(mutation\_locations))  
{  
 mutation\_location\_location <- which(mutation\_locations[i,] != "NA")  
 if(length(mutation\_location\_location) > 0)  
 {  
 for(j in 1:length(mutation\_location\_location))  
 {  
 print("-----start------")  
 print("i=")  
 print(i)  
 print("j=")  
 print(j)  
 print("mutation\_locations[i,j]=")  
 print(mutation\_locations[i,j])  
 mutation\_location\_x <- as.numeric(mutation\_locations[i,j])   
 print("mutation\_location\_x=")  
 print(mutation\_location\_x)  
 mutation\_location\_remove\_dash <- mutation\_location\_x - sum(trim\_alignment[1,1:(mutation\_location\_x-1)] == "-")  
 print("mutation\_location\_remove\_dash=")  
 print(mutation\_location\_remove\_dash)  
 modulo\_mut\_loc <- mutation\_location\_remove\_dash %% 3  
 modulo\_mut\_loc\_df[i,j] <- modulo\_mut\_loc  
   
 if(trim\_alignment[1,mutation\_location\_x] == "-")  
 {  
 #print("plop")  
 AAmutations\_df[i,j] <- "frame\_shift\_ins"  
 #print(AAmutations\_df[i,j])  
 }   
   
 else if (modulo\_mut\_loc == 0)  
 {  
 codon\_start <- mutation\_location\_remove\_dash-2  
 codon\_end <- mutation\_location\_remove\_dash  
 }  
 else if (modulo\_mut\_loc == 1)  
 {  
 codon\_start <- mutation\_location\_remove\_dash  
 codon\_end <- mutation\_location\_remove\_dash+2  
 }  
 else if (modulo\_mut\_loc == 2)  
 {  
 codon\_start <- mutation\_location\_remove\_dash-1  
 codon\_end <- mutation\_location\_remove\_dash+1  
 }  
   
 aa\_position <- ceiling(mutation\_location\_remove\_dash/3)  
 print("aaPosition")  
 print(aa\_position)  
   
 WTcodon <- paste(c(trim\_dash[1,codon\_start:codon\_end]), collapse="")  
 WTcodon\_DNAstr <- DNAString(WTcodon)  
 WTaa <- translate(WTcodon\_DNAstr)  
 print("WTcodon and aa")  
 print(WTcodon\_DNAstr)  
 print(WTaa)  
   
 MUTcodon <- c(trim\_dash[i,codon\_start:codon\_end])  
 print("MUTcodon")  
 print(MUTcodon)  
   
 if (sum(MUTcodon[]== "-") != 0)  
 {  
 MUTaa <- "frame\_shift\_del"  
 print("frame\_shift\_del")  
 }  
   
   
 else if (sum(MUTcodon[] %in% c("A","T","C","G")) == 3)  
 {  
 MUTcodon\_DNAstr <- DNAString(paste(MUTcodon, collapse = ""))  
 print("MUTcodon\_DNAstr")  
 print(MUTcodon\_DNAstr)  
 MUTaa <- translate(MUTcodon\_DNAstr)  
 print("translation!MUTaa")  
 print(MUTaa)  
 }  
   
 else if (sum(MUTcodon[] %in% c("A","T","C","G","-")) != 3)  
 {  
 MUTaa <- "unidentifiable base"  
 print("unidentifiable base :(")  
 }  
   
 WTaa\_cha <- as.character(WTaa)  
 MUTaa\_cha <- as.character(MUTaa)  
 aa\_position\_cha <- as.character(aa\_position)  
 AAmutations <- c(WTaa\_cha,aa\_position\_cha,MUTaa\_cha)  
 AAmutations\_df[i,j] <- paste(AAmutations, collapse="")  
 print("-----END-----")  
 }  
 }  
}

## [1] "-----start------"  
## [1] "i="  
## [1] 2  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 11\_T7 Biotec\_CGN   
## 181   
## [1] "mutation\_location\_x="  
## [1] 181  
## [1] "mutation\_location\_remove\_dash="  
## [1] 181  
## [1] "aaPosition"  
## [1] 61  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AAT  
## 1-letter AAString object  
## seq: N  
## [1] "MUTcodon"  
## $V237  
## [1] "G"  
##   
## $V238  
## [1] "A"  
##   
## $V239  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GAT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: D  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 8  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 22\_T7 Biotec\_CGN   
## 741   
## [1] "mutation\_location\_x="  
## [1] 741  
## [1] "mutation\_location\_remove\_dash="  
## [1] 741  
## [1] "aaPosition"  
## [1] 247  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AAT  
## 1-letter AAString object  
## seq: N  
## [1] "MUTcodon"  
## $V237  
## [1] "A"  
##   
## $V238  
## [1] "A"  
##   
## $V239  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: AAT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: N  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 9  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 23\_T7 Biotec\_CGN   
## 273   
## [1] "mutation\_location\_x="  
## [1] 273  
## [1] "mutation\_location\_remove\_dash="  
## [1] 273  
## [1] "aaPosition"  
## [1] 91  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GCA  
## 1-letter AAString object  
## seq: A  
## [1] "MUTcodon"  
## $V327  
## [1] "G"  
##   
## $V328  
## [1] "C"  
##   
## $V329  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GCG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: A  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 11  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 25\_T7 Biotec\_CGN   
## 593   
## [1] "mutation\_location\_x="  
## [1] 593  
## [1] "mutation\_location\_remove\_dash="  
## [1] 593  
## [1] "aaPosition"  
## [1] 198  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TAT  
## 1-letter AAString object  
## seq: Y  
## [1] "MUTcodon"  
## $V648  
## [1] "T"  
##   
## $V649  
## [1] "T"  
##   
## $V650  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: TTT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: F  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 12  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 31\_T7 Biotec\_CGN   
## 104   
## [1] "mutation\_location\_x="  
## [1] 104  
## [1] "mutation\_location\_remove\_dash="  
## [1] 104  
## [1] "aaPosition"  
## [1] 35  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GAC  
## 1-letter AAString object  
## seq: D  
## [1] "MUTcodon"  
## $V159  
## [1] "G"  
##   
## $V160  
## [1] "G"  
##   
## $V161  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: G  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 12  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 31\_T7 Biotec\_CGN   
## 628   
## [1] "mutation\_location\_x="  
## [1] 628  
## [1] "mutation\_location\_remove\_dash="  
## [1] 628  
## [1] "aaPosition"  
## [1] 210  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: ATC  
## 1-letter AAString object  
## seq: I  
## [1] "MUTcodon"  
## $V684  
## [1] "G"  
##   
## $V685  
## [1] "T"  
##   
## $V686  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GTC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: V  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 13  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 32\_T7 Biotec\_CGN   
## 169   
## [1] "mutation\_location\_x="  
## [1] 169  
## [1] "mutation\_location\_remove\_dash="  
## [1] 169  
## [1] "aaPosition"  
## [1] 57  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AAG  
## 1-letter AAString object  
## seq: K  
## [1] "MUTcodon"  
## $V225  
## [1] "G"  
##   
## $V226  
## [1] "A"  
##   
## $V227  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GAG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: E  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 13  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 32\_T7 Biotec\_CGN   
## 186   
## [1] "mutation\_location\_x="  
## [1] 186  
## [1] "mutation\_location\_remove\_dash="  
## [1] 186  
## [1] "aaPosition"  
## [1] 62  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GCT  
## 1-letter AAString object  
## seq: A  
## [1] "MUTcodon"  
## $V240  
## [1] "G"  
##   
## $V241  
## [1] "C"  
##   
## $V242  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GCC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: A  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 13  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 32\_T7 Biotec\_CGN   
## 253   
## [1] "mutation\_location\_x="  
## [1] 253  
## [1] "mutation\_location\_remove\_dash="  
## [1] 253  
## [1] "aaPosition"  
## [1] 85  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CCA  
## 1-letter AAString object  
## seq: P  
## [1] "MUTcodon"  
## $V309  
## [1] "T"  
##   
## $V310  
## [1] "C"  
##   
## $V311  
## [1] "A"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: TCA  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: S  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 13  
## [1] "j="  
## [1] 4  
## [1] "mutation\_locations[i,j]="  
## 32\_T7 Biotec\_CGN   
## 365   
## [1] "mutation\_location\_x="  
## [1] 365  
## [1] "mutation\_location\_remove\_dash="  
## [1] 365  
## [1] "aaPosition"  
## [1] 122  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GAC  
## 1-letter AAString object  
## seq: D  
## [1] "MUTcodon"  
## $V420  
## [1] "G"  
##   
## $V421  
## [1] "G"  
##   
## $V422  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: G  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 14  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 33\_T7 Biotec\_CGN   
## 15   
## [1] "mutation\_location\_x="  
## [1] 15  
## [1] "mutation\_location\_remove\_dash="  
## [1] 15  
## [1] "aaPosition"  
## [1] 5  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CAT  
## 1-letter AAString object  
## seq: H  
## [1] "MUTcodon"  
## $V69  
## [1] "C"  
##   
## $V70  
## [1] "A"  
##   
## $V71  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CAC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: H  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 16  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 35\_T7 Biotec\_CGN   
## 149   
## [1] "mutation\_location\_x="  
## [1] 149  
## [1] "mutation\_location\_remove\_dash="  
## [1] 149  
## [1] "aaPosition"  
## [1] 50  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TTT  
## 1-letter AAString object  
## seq: F  
## [1] "MUTcodon"  
## $V204  
## [1] "T"  
##   
## $V205  
## [1] "C"  
##   
## $V206  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: TCT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: S  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 17  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 41\_T7 Biotec\_CGN   
## 29   
## [1] "mutation\_location\_x="  
## [1] 29  
## [1] "mutation\_location\_remove\_dash="  
## [1] 29  
## [1] "aaPosition"  
## [1] 10  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CCG  
## 1-letter AAString object  
## seq: P  
## [1] "MUTcodon"  
## $V84  
## [1] "C"  
##   
## $V85  
## [1] "T"  
##   
## $V86  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CTG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: M  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 18  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 42\_T7 Biotec\_CGN   
## 90   
## [1] "mutation\_location\_x="  
## [1] 90  
## [1] "mutation\_location\_remove\_dash="  
## [1] 90  
## [1] "aaPosition"  
## [1] 30  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TCC  
## 1-letter AAString object  
## seq: S  
## [1] "MUTcodon"  
## $V144  
## [1] "T"  
##   
## $V145  
## [1] "C"  
##   
## $V146  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: TCT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: S  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 18  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 42\_T7 Biotec\_CGN   
## 501   
## [1] "mutation\_location\_x="  
## [1] 501  
## [1] "mutation\_location\_remove\_dash="  
## [1] 501  
## [1] "aaPosition"  
## [1] 167  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CGT  
## 1-letter AAString object  
## seq: R  
## [1] "MUTcodon"  
## $V555  
## [1] "C"  
##   
## $V556  
## [1] "G"  
##   
## $V557  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: R  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 18  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 42\_T7 Biotec\_CGN   
## 543   
## [1] "mutation\_location\_x="  
## [1] 543  
## [1] "mutation\_location\_remove\_dash="  
## [1] 543  
## [1] "aaPosition"  
## [1] 181  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GCG  
## 1-letter AAString object  
## seq: A  
## [1] "MUTcodon"  
## $V597  
## [1] "G"  
##   
## $V598  
## [1] "C"  
##   
## $V599  
## [1] "A"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GCA  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: A  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 19  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 43\_T7 Biotec\_CGN   
## 108   
## [1] "mutation\_location\_x="  
## [1] 108  
## [1] "mutation\_location\_remove\_dash="  
## [1] 108  
## [1] "aaPosition"  
## [1] 36  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GGA  
## 1-letter AAString object  
## seq: G  
## [1] "MUTcodon"  
## $V162  
## [1] "G"  
##   
## $V163  
## [1] "G"  
##   
## $V164  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GGG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: G  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 20  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 44\_T7 Biotec\_CGN   
## 117   
## [1] "mutation\_location\_x="  
## [1] 117  
## [1] "mutation\_location\_remove\_dash="  
## [1] 117  
## [1] "aaPosition"  
## [1] 39  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GGT  
## 1-letter AAString object  
## seq: G  
## [1] "MUTcodon"  
## $V171  
## [1] "G"  
##   
## $V172  
## [1] "G"  
##   
## $V173  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: G  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 20  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 44\_T7 Biotec\_CGN   
## 126   
## [1] "mutation\_location\_x="  
## [1] 126  
## [1] "mutation\_location\_remove\_dash="  
## [1] 126  
## [1] "aaPosition"  
## [1] 42  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TGT  
## 1-letter AAString object  
## seq: C  
## [1] "MUTcodon"  
## $V180  
## [1] "T"  
##   
## $V181  
## [1] "G"  
##   
## $V182  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: TGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: C  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 20  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 44\_T7 Biotec\_CGN   
## 251   
## [1] "mutation\_location\_x="  
## [1] 251  
## [1] "mutation\_location\_remove\_dash="  
## [1] 251  
## [1] "aaPosition"  
## [1] 84  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GTA  
## 1-letter AAString object  
## seq: V  
## [1] "MUTcodon"  
## $V306  
## [1] "G"  
##   
## $V307  
## [1] "A"  
##   
## $V308  
## [1] "A"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GAA  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: E  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 21  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 45\_T7 Biotec\_CGN   
## 53   
## [1] "mutation\_location\_x="  
## [1] 53  
## [1] "mutation\_location\_remove\_dash="  
## [1] 53  
## [1] "aaPosition"  
## [1] 18  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AAG  
## 1-letter AAString object  
## seq: K  
## [1] "MUTcodon"  
## $V108  
## [1] "A"  
##   
## $V109  
## [1] "G"  
##   
## $V110  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: AGG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: R  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 21  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 45\_T7 Biotec\_CGN   
## 310   
## [1] "mutation\_location\_x="  
## [1] 310  
## [1] "mutation\_location\_remove\_dash="  
## [1] 310  
## [1] "aaPosition"  
## [1] 104  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GGG  
## 1-letter AAString object  
## seq: G  
## [1] "MUTcodon"  
## $V366  
## [1] "A"  
##   
## $V367  
## [1] "G"  
##   
## $V368  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: AGG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: R  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 22  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 46\_T7 Biotec\_CGN   
## 119   
## [1] "mutation\_location\_x="  
## [1] 119  
## [1] "mutation\_location\_remove\_dash="  
## [1] 119  
## [1] "aaPosition"  
## [1] 40  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CGT  
## 1-letter AAString object  
## seq: R  
## [1] "MUTcodon"  
## $V174  
## [1] "C"  
##   
## $V175  
## [1] "A"  
##   
## $V176  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CAT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: H  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 22  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 46\_T7 Biotec\_CGN   
## 169   
## [1] "mutation\_location\_x="  
## [1] 169  
## [1] "mutation\_location\_remove\_dash="  
## [1] 169  
## [1] "aaPosition"  
## [1] 57  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AAG  
## 1-letter AAString object  
## seq: K  
## [1] "MUTcodon"  
## $V225  
## [1] "G"  
##   
## $V226  
## [1] "A"  
##   
## $V227  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GAG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: E  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 22  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 46\_T7 Biotec\_CGN   
## 384   
## [1] "mutation\_location\_x="  
## [1] 384  
## [1] "mutation\_location\_remove\_dash="  
## [1] 384  
## [1] "aaPosition"  
## [1] 128  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AGT  
## 1-letter AAString object  
## seq: S  
## [1] "MUTcodon"  
## $V438  
## [1] "A"  
##   
## $V439  
## [1] "G"  
##   
## $V440  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: AGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: S  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 23  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 51\_T7 Biotec\_CGN   
## 124   
## [1] "mutation\_location\_x="  
## [1] 124  
## [1] "mutation\_location\_remove\_dash="  
## [1] 124  
## [1] "aaPosition"  
## [1] 42  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TGT  
## 1-letter AAString object  
## seq: C  
## [1] "MUTcodon"  
## $V180  
## [1] "A"  
##   
## $V181  
## [1] "G"  
##   
## $V182  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: AGT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: S  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 23  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 51\_T7 Biotec\_CGN   
## 184   
## [1] "mutation\_location\_x="  
## [1] 184  
## [1] "mutation\_location\_remove\_dash="  
## [1] 184  
## [1] "aaPosition"  
## [1] 62  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GCT  
## 1-letter AAString object  
## seq: A  
## [1] "MUTcodon"  
## $V240  
## [1] "A"  
##   
## $V241  
## [1] "C"  
##   
## $V242  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: ACT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: T  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 23  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 51\_T7 Biotec\_CGN   
## 241   
## [1] "mutation\_location\_x="  
## [1] 241  
## [1] "mutation\_location\_remove\_dash="  
## [1] 241  
## [1] "aaPosition"  
## [1] 81  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AGC  
## 1-letter AAString object  
## seq: S  
## [1] "MUTcodon"  
## $V297  
## [1] "G"  
##   
## $V298  
## [1] "G"  
##   
## $V299  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: G  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 23  
## [1] "j="  
## [1] 4  
## [1] "mutation\_locations[i,j]="  
## 51\_T7 Biotec\_CGN   
## 322   
## [1] "mutation\_location\_x="  
## [1] 322  
## [1] "mutation\_location\_remove\_dash="  
## [1] 322  
## [1] "aaPosition"  
## [1] 108  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: ATC  
## 1-letter AAString object  
## seq: I  
## [1] "MUTcodon"  
## $V378  
## [1] "-"  
##   
## $V379  
## [1] "T"  
##   
## $V380  
## [1] "C"  
##   
## [1] "frame\_shift\_del"  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 23  
## [1] "j="  
## [1] 5  
## [1] "mutation\_locations[i,j]="  
## 51\_T7 Biotec\_CGN   
## 594   
## [1] "mutation\_location\_x="  
## [1] 594  
## [1] "mutation\_location\_remove\_dash="  
## [1] 594  
## [1] "aaPosition"  
## [1] 198  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TAT  
## 1-letter AAString object  
## seq: Y  
## [1] "MUTcodon"  
## $V648  
## [1] "T"  
##   
## $V649  
## [1] "A"  
##   
## $V650  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: TAC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: Y  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 23  
## [1] "j="  
## [1] 6  
## [1] "mutation\_locations[i,j]="  
## 51\_T7 Biotec\_CGN   
## 679   
## [1] "mutation\_location\_x="  
## [1] 679  
## [1] "mutation\_location\_remove\_dash="  
## [1] 679  
## [1] "aaPosition"  
## [1] 227  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: AAG  
## 1-letter AAString object  
## seq: K  
## [1] "MUTcodon"  
## $V735  
## [1] "G"  
##   
## $V736  
## [1] "A"  
##   
## $V737  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GAG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: E  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 25  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 53\_T7 Biotec\_CGN   
## 29   
## [1] "mutation\_location\_x="  
## [1] 29  
## [1] "mutation\_location\_remove\_dash="  
## [1] 29  
## [1] "aaPosition"  
## [1] 10  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CCG  
## 1-letter AAString object  
## seq: P  
## [1] "MUTcodon"  
## $V84  
## [1] "C"  
##   
## $V85  
## [1] "T"  
##   
## $V86  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CTG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: M  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 25  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 53\_T7 Biotec\_CGN   
## 456   
## [1] "mutation\_location\_x="  
## [1] 456  
## [1] "mutation\_location\_remove\_dash="  
## [1] 456  
## [1] "aaPosition"  
## [1] 152  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CCC  
## 1-letter AAString object  
## seq: P  
## [1] "MUTcodon"  
## $V510  
## [1] "C"  
##   
## $V511  
## [1] "C"  
##   
## $V512  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CCT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: P  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 25  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 53\_T7 Biotec\_CGN   
## 460   
## [1] "mutation\_location\_x="  
## [1] 460  
## [1] "mutation\_location\_remove\_dash="  
## [1] 460  
## [1] "aaPosition"  
## [1] 154  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TAC  
## 1-letter AAString object  
## seq: Y  
## [1] "MUTcodon"  
## $V516  
## [1] "C"  
##   
## $V517  
## [1] "A"  
##   
## $V518  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CAC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: H  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 25  
## [1] "j="  
## [1] 4  
## [1] "mutation\_locations[i,j]="  
## 53\_T7 Biotec\_CGN   
## 579   
## [1] "mutation\_location\_x="  
## [1] 579  
## [1] "mutation\_location\_remove\_dash="  
## [1] 579  
## [1] "aaPosition"  
## [1] 193  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GAA  
## 1-letter AAString object  
## seq: E  
## [1] "MUTcodon"  
## $V633  
## [1] "G"  
##   
## $V634  
## [1] "A"  
##   
## $V635  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GAG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: E  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 26  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 54\_T7 Biotec\_CGN   
## 177   
## [1] "mutation\_location\_x="  
## [1] 177  
## [1] "mutation\_location\_remove\_dash="  
## [1] 177  
## [1] "aaPosition"  
## [1] 59  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: CCA  
## 1-letter AAString object  
## seq: P  
## [1] "MUTcodon"  
## $V231  
## [1] "C"  
##   
## $V232  
## [1] "C"  
##   
## $V233  
## [1] "G"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CCG  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: P  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 26  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 54\_T7 Biotec\_CGN   
## 326   
## [1] "mutation\_location\_x="  
## [1] 326  
## [1] "mutation\_location\_remove\_dash="  
## [1] 326  
## [1] "aaPosition"  
## [1] 109  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GTT  
## 1-letter AAString object  
## seq: V  
## [1] "MUTcodon"  
## $V381  
## [1] "G"  
##   
## $V382  
## [1] "C"  
##   
## $V383  
## [1] "T"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GCT  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: A  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 26  
## [1] "j="  
## [1] 3  
## [1] "mutation\_locations[i,j]="  
## 54\_T7 Biotec\_CGN   
## 603   
## [1] "mutation\_location\_x="  
## [1] 603  
## [1] "mutation\_location\_remove\_dash="  
## [1] 603  
## [1] "aaPosition"  
## [1] 201  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GGT  
## 1-letter AAString object  
## seq: G  
## [1] "MUTcodon"  
## $V657  
## [1] "G"  
##   
## $V658  
## [1] "G"  
##   
## $V659  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GGC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: G  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 27  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 55\_T7 Biotec\_CGN   
## 229   
## [1] "mutation\_location\_x="  
## [1] 229  
## [1] "mutation\_location\_remove\_dash="  
## [1] 229  
## [1] "aaPosition"  
## [1] 77  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TTC  
## 1-letter AAString object  
## seq: F  
## [1] "MUTcodon"  
## $V285  
## [1] "C"  
##   
## $V286  
## [1] "T"  
##   
## $V287  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CTC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: L  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 27  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 55\_T7 Biotec\_CGN   
## 327   
## [1] "mutation\_location\_x="  
## [1] 327  
## [1] "mutation\_location\_remove\_dash="  
## [1] 327  
## [1] "aaPosition"  
## [1] 109  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GTT  
## 1-letter AAString object  
## seq: V  
## [1] "MUTcodon"  
## $V381  
## [1] "G"  
##   
## $V382  
## [1] "T"  
##   
## $V383  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GTC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: V  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 28  
## [1] "j="  
## [1] 1  
## [1] "mutation\_locations[i,j]="  
## 56\_T7 Biotec\_CGN   
## 229   
## [1] "mutation\_location\_x="  
## [1] 229  
## [1] "mutation\_location\_remove\_dash="  
## [1] 229  
## [1] "aaPosition"  
## [1] 77  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: TTC  
## 1-letter AAString object  
## seq: F  
## [1] "MUTcodon"  
## $V285  
## [1] "C"  
##   
## $V286  
## [1] "T"  
##   
## $V287  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: CTC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: L  
## [1] "-----END-----"  
## [1] "-----start------"  
## [1] "i="  
## [1] 28  
## [1] "j="  
## [1] 2  
## [1] "mutation\_locations[i,j]="  
## 56\_T7 Biotec\_CGN   
## 327   
## [1] "mutation\_location\_x="  
## [1] 327  
## [1] "mutation\_location\_remove\_dash="  
## [1] 327  
## [1] "aaPosition"  
## [1] 109  
## [1] "WTcodon and aa"  
## 3-letter DNAString object  
## seq: GTT  
## 1-letter AAString object  
## seq: V  
## [1] "MUTcodon"  
## $V381  
## [1] "G"  
##   
## $V382  
## [1] "T"  
##   
## $V383  
## [1] "C"  
##   
## [1] "MUTcodon\_DNAstr"  
## 3-letter DNAString object  
## seq: GTC  
## [1] "translation!MUTaa"  
## 1-letter AAString object  
## seq: V  
## [1] "-----END-----"

show results: AAmutations\_df

AAmutations\_df

## [,1] [,2] [,3] [,4] [,5] [,6]   
## ClyF NA NA NA NA NA NA   
## 11\_T7 Biotec\_CGN "N61D" NA NA NA NA NA   
## 12\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 13\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 14\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 15\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 21\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 22\_T7 Biotec\_CGN "N247N" NA NA NA NA NA   
## 23\_T7 Biotec\_CGN "A91A" NA NA NA NA NA   
## 24\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 25\_T7 Biotec\_CGN "Y198F" NA NA NA NA NA   
## 31\_T7 Biotec\_CGN "D35G" "I210V" NA NA NA NA   
## 32\_T7 Biotec\_CGN "K57E" "A62A" "P85S" "D122G" NA NA   
## 33\_T7 Biotec\_CGN "H5H" NA NA NA NA NA   
## 34\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 35\_T7 Biotec\_CGN "F50S" NA NA NA NA NA   
## 41\_T7 Biotec\_CGN "P10M" NA NA NA NA NA   
## 42\_T7 Biotec\_CGN "S30S" "R167R" "A181A" NA NA NA   
## 43\_T7 Biotec\_CGN "G36G" NA NA NA NA NA   
## 44\_T7 Biotec\_CGN "G39G" "C42C" "V84E" NA NA NA   
## 45\_T7 Biotec\_CGN "K18R" "G104R" NA NA NA NA   
## 46\_T7 Biotec\_CGN "R40H" "K57E" "S128S" NA NA NA   
## 51\_T7 Biotec\_CGN "C42S" "A62T" "S81G" "I108frame\_shift\_del" "Y198Y" "K227E"  
## 52\_T7 Biotec\_CGN NA NA NA NA NA NA   
## 53\_T7 Biotec\_CGN "P10M" "P152P" "Y154H" "E193E" NA NA   
## 54\_T7 Biotec\_CGN "P59P" "V109A" "G201G" NA NA NA   
## 55\_T7 Biotec\_CGN "F77L" "V109V" NA NA NA NA   
## 56\_T7 Biotec\_CGN "F77L" "V109V" NA NA NA NA

Show results: DNA mutation location (alignment location)

mutation\_locations

## [,1] [,2] [,3] [,4] [,5] [,6]  
## ClyF NA NA NA NA NA NA  
## 11\_T7 Biotec\_CGN 181 NA NA NA NA NA  
## 12\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 13\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 14\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 15\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 21\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 22\_T7 Biotec\_CGN 741 NA NA NA NA NA  
## 23\_T7 Biotec\_CGN 273 NA NA NA NA NA  
## 24\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 25\_T7 Biotec\_CGN 593 NA NA NA NA NA  
## 31\_T7 Biotec\_CGN 104 628 NA NA NA NA  
## 32\_T7 Biotec\_CGN 169 186 253 365 NA NA  
## 33\_T7 Biotec\_CGN 15 NA NA NA NA NA  
## 34\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 35\_T7 Biotec\_CGN 149 NA NA NA NA NA  
## 41\_T7 Biotec\_CGN 29 NA NA NA NA NA  
## 42\_T7 Biotec\_CGN 90 501 543 NA NA NA  
## 43\_T7 Biotec\_CGN 108 NA NA NA NA NA  
## 44\_T7 Biotec\_CGN 117 126 251 NA NA NA  
## 45\_T7 Biotec\_CGN 53 310 NA NA NA NA  
## 46\_T7 Biotec\_CGN 119 169 384 NA NA NA  
## 51\_T7 Biotec\_CGN 124 184 241 322 594 679  
## 52\_T7 Biotec\_CGN NA NA NA NA NA NA  
## 53\_T7 Biotec\_CGN 29 456 460 579 NA NA  
## 54\_T7 Biotec\_CGN 177 326 603 NA NA NA  
## 55\_T7 Biotec\_CGN 229 327 NA NA NA NA  
## 56\_T7 Biotec\_CGN 229 327 NA NA NA NA

Show results: DNA mutation counts. -: no mutation AG: A to G GA: G to A Del: Deletion Ins: Insertion ?: Not recognized nucleotide

mutation\_count\_row

## - AG GA AT TA AC CA GC CG GT TG CT TC Del Ins ?  
## ClyF 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 11\_T7 Biotec\_CGN 753 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 12\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 13\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 14\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 15\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 21\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 22\_T7 Biotec\_CGN 753 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0  
## 23\_T7 Biotec\_CGN 753 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 24\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 25\_T7 Biotec\_CGN 753 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 31\_T7 Biotec\_CGN 752 2 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 32\_T7 Biotec\_CGN 750 2 0 0 0 0 0 0 0 0 0 1 1 0 0 0  
## 33\_T7 Biotec\_CGN 753 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 34\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 35\_T7 Biotec\_CGN 753 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 41\_T7 Biotec\_CGN 753 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0  
## 42\_T7 Biotec\_CGN 751 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0  
## 43\_T7 Biotec\_CGN 753 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 44\_T7 Biotec\_CGN 751 0 0 0 1 0 0 0 0 0 0 0 2 0 0 0  
## 45\_T7 Biotec\_CGN 752 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 46\_T7 Biotec\_CGN 751 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0  
## 51\_T7 Biotec\_CGN 748 2 1 0 1 0 0 0 0 0 0 0 1 1 0 0  
## 52\_T7 Biotec\_CGN 754 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 53\_T7 Biotec\_CGN 750 1 0 0 0 0 0 0 0 0 0 2 1 0 0 0  
## 54\_T7 Biotec\_CGN 751 1 0 0 0 0 0 0 0 0 0 0 2 0 0 0  
## 55\_T7 Biotec\_CGN 752 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0  
## 56\_T7 Biotec\_CGN 752 0 0 0 0 0 0 0 0 0 0 0 2 0 0 0

export the result tables as csv

##turn "NA" to nothing for output csv  
AAmutations\_removeNA <- AAmutations\_df  
AAmutations\_removeNA[is.na(AAmutations\_removeNA)] <- ""  
mutation\_locations\_removeNA <- mutation\_locations  
mutation\_locations\_removeNA[is.na(mutation\_locations\_removeNA)] <- ""  
##create a new folder in the current working directory   
dir.create("mutation\_analysis\_results", recursive = T)

## Warning in dir.create("mutation\_analysis\_results", recursive = T):  
## 'mutation\_analysis\_results' already exists

##Save the dataframes   
write.csv(mutation\_count\_row,"mutation\_analysis\_results/DNAmutation\_counts.csv", row.names = TRUE)  
write.csv(mutation\_locations\_removeNA,"mutation\_analysis\_results/DNAmutation\_locations.csv", row.names = TRUE)  
write.csv(AAmutations\_removeNA,"mutation\_analysis\_results/AAmutations.csv", row.names = TRUE)  
  
list.files('mutation\_analysis\_results',full.names = T)

## [1] "mutation\_analysis\_results/AAmutations.csv"   
## [2] "mutation\_analysis\_results/DNAmutation\_counts.csv"   
## [3] "mutation\_analysis\_results/DNAmutation\_locations.csv"