

Clustering Polypeptide motifs using the Hammock hidden Markov model peptide clustering of organoid samples

Tomas Bjorklund

Mon Nov 9 22:42:05 2020

This script clusters Polypeptide motifs using the Hammock hidden Markov model peptide clustering fo Organoid samples.

```
suppressPackageStartupMessages(library(knitr))
```

Loading samples

```
all.samples <- readRDS("data/allSamplesDataTable.RDS")
all.samples[, `:=`(Peptide, as.character(Peptide)), ]

setkey(all.samples, Group)

select.samples <- all.samples[J(c("mRNA_30cpc_Organoid_MD114", "mRNA_3000cpc_Organoid_MD101"))]

select.samples[, `:=`(BCcount, as.integer(mclapply(BC, function(x) length(table(strsplit(paste(t(x),
collapse = ","), ","))), mc.cores = detectCores())))]
select.samples[, `:=`(Animalcount, as.integer(mclapply(Animals, function(x) length(table(strsplit(paste(t(x),
collapse = ","), ","))), mc.cores = detectCores())))]
select.samples[Animals == "mRNA_3000cpc_Organoid_MD101", `:=`(BCcount, as.integer(BCcount/5)),
] # Removes unclear 3000cpc reads
select.samples <- select.samples[BCcount >= 1]
select.samples[, `:=`(Score, BCcount + Animalcount - 1), ]
select.samples.trsp <- unique(select.samples, by = c("Animals", "BC", "LUTnrs"))

fasta.names <- paste(1:nrow(select.samples.trsp), select.samples.trsp$Score,
select.samples.trsp$Group, sep = "|")
write.fasta(as.list(select.samples.trsp$Peptide), fasta.names, "data/organoidPeptides.fasta",
open = "w", nbchar = 60, as.string = TRUE)

# Generate Scoring table for Weblogo Weighting
select.samples.pepMerge <- select.samples.trsp[, sum(Score), by = c("Peptide")]
setnames(select.samples.pepMerge, "V1", "Score")
```

Executing Hammock Clustering

```
Sys.setenv(PATH = paste("/root/HMMER/binaries", Sys.getenv("PATH"), sep = ":"),
HHLIB = "/home/rstudio/Hammock_v_1.2.0/hhsuite-2.0.16/lib/hh/")
unlink("/home/rstudio/data/HammockOrganoids", recursive = TRUE, force = FALSE)
sys.out <- system(paste("java -jar /home/rstudio/Hammock_v_1.2.0/dist/Hammock.jar full -i /home/rstudio/data/
detectCores(), sep = ""), intern = TRUE, ignore.stdout = TRUE)
# Alternative parameters --use_clinkage --alignment_threshold 23 --max_shift
# 13 --max_aln_length 37 --count_threshold 50 --max_inner_gaps 0
# --assign_thresholds 14.1,10.5,7.0
hammock.log <- data.table(readLines("data/HammockOrganoids/run.log"))
```

```
colnames(hammock.log) <- c("Hammock log file")
knitr::kable(hammock.log, longtable = T)
```

Hammock log file

2020-11-09 22:42:17.002:

Hammock version 1.2.0 Run with `-help` for a brief description of command line parameters.

2020-11-09 22:42:17.136: Program started in mode “full”.

Command-line arguments:

full -i /home/rstudio/data/organoidPeptides.fasta -d /home/rstudio/data/HammockOrganoids -max_shift 7 -c 65 -t 48

Complete list of input/output parameters:

-i, -input /home/rstudio/data/organoidPeptides.fasta
-d, -output_directory /home/rstudio/data/HammockOrganoids
-t, -thread 48
-l, -labels null

Complete list of linkage clustering parameters:

-f, -file_format fasta
-m, -matrix /home/rstudio/Hammock_v_1.2.0/matrices/blosum62.txt
-g, -alignment_threshold (-greedy_threshold)null
-x, -max_shift 7
-p, -gap_penalty 0
-C, -cache_size_limit 1

2020-11-09 22:42:17.137: Loading input sequences...

2020-11-09 22:42:17.173: 1934 unique sequences loaded.

2020-11-09 22:42:17.176: 2141 total sequences loaded.

2020-11-09 22:42:17.176: 1934 unique sequences after non-specified labels filtered out

2020-11-09 22:42:17.181: 2141 total sequences after non-specified labels filtered out

2020-11-09 22:42:17.181: Shortest sequence: 14 AA. Longest sequence: 22 AA.

2020-11-09 22:42:17.182: Up to 10 000 unique sequences. Using linkage clustering. Use `-use_greedy` to force greedy clustering

2020-11-09 22:42:17.187: Generating input statistics...

2020-11-09 22:42:17.194: Linkage clustering threshold not set. Setting automatically to: 26

2020-11-09 22:42:17.195: Linkage clustering...

2020-11-09 22:42:25.699: Ready. Clustering time: 8503

2020-11-09 22:42:25.700: Resulting clusters: 1024

2020-11-09 22:42:25.700: Building MSAs...

2020-11-09 22:42:26.086: Ready. Total time: 8890

2020-11-09 22:42:26.086: Saving results to output files...

2020-11-09 22:42:26.217: Linkage clustering results in: /home/rstudio/data/HammockOrganoids/initial_clusters.tsv

2020-11-09 22:42:26.217: and: /home/rstudio/data/HammockOrganoids/initial_clusters_sequences.tsv

2020-11-09 22:42:26.218: and:

/home/rstudio/data/HammockOrganoids/initial_clusters_sequences_original_order.tsv

2020-11-09 22:42:26.218:

Loading clusters...

2020-11-09 22:42:26.238: Maximal alignment length not set. Setting automatically to: 30

2020-11-09 22:42:26.239: Minimal number of match states not set. Setting automatically to: 5

2020-11-09 22:42:26.298: Assign threshold not set. Setting automatically to:

2020-11-09 22:42:26.302: 14.36,11.34,8.31

2020-11-09 22:42:26.302: Overlap threshold not set. Setting automatically to:

2020-11-09 22:42:26.303: 10.58,6.05,0.0

Hammock log file

2020-11-09 22:42:26.303: Merge threshold not set. Setting automatically on the basis of average sequence length to:
2020-11-09 22:42:26.304: 15.12,13.61,12.09
2020-11-09 22:42:26.305: Initial extension threshold not set. Setting automatically on the basis of merge threshold
sequence to: 16.632

Complete list of HMM-based clustering parameters:

-a, -part_threshold null
-s, -size_threshold null
-c, -count_threshold 65
-n, -assign_thresholds 14.36,11.34,8.31,
-v, -overlap_thresholds 10.58,6.05,0.0,
-r, -merge_thresholds 15.12,13.61,12.09,
-e, -relative_thresholds false
-b, -absolute_thresholds true
-h, -min_conserved_positions 5
-y, -max_gap_proportion 0.2
-k, -min_ic 1.2
-j, -max_aln_length 30
-u, -max_inner_gaps 0
-q, -extension_increase_length false

2020-11-09 22:42:26.385: Initial cluster extension...
2020-11-09 22:42:51.575:
Clustering in 3 rounds...
2020-11-09 22:42:51.575:
2020-11-09 22:42:51.576: Round 1:

2020-11-09 22:42:51.576: 65 clusters remaining
2020-11-09 22:42:51.577: Building hmms and searching database...
2020-11-09 22:42:51.818: Extending clusters...
2020-11-09 22:42:51.820: 13 sequences to be inserted into clusters
2020-11-09 22:42:51.820: 12 clusters to be extended
2020-11-09 22:42:51.829: 11 sequences rejected
2020-11-09 22:42:51.830: 1 cluster pairs to check and merge.
2020-11-09 22:42:51.830: Merging clusters from 1 groups...
2020-11-09 22:42:51.838: Buiding hhs...
2020-11-09 22:42:51.839: HH clustering...
2020-11-09 22:42:52.108:
2020-11-09 22:42:52.108: Round 2:

2020-11-09 22:42:52.109: 65 clusters remaining
2020-11-09 22:42:52.109: Building hmms and searching database...
2020-11-09 22:42:52.184: Extending clusters...
2020-11-09 22:42:52.186: 20 sequences to be inserted into clusters
2020-11-09 22:42:52.186: 17 clusters to be extended
2020-11-09 22:42:52.196: 19 sequences rejected
2020-11-09 22:42:52.197: 24 cluster pairs to check and merge.
2020-11-09 22:42:52.197: Merging clusters from 6 groups...
2020-11-09 22:42:52.206: Buiding hhs...
2020-11-09 22:42:52.218: HH clustering...
2020-11-09 22:42:53.356:
2020-11-09 22:42:53.356: Round 3:

2020-11-09 22:42:53.356: 62 clusters remaining
2020-11-09 22:42:53.356: Building hmms and searching database...
2020-11-09 22:42:53.450: Extending clusters...

Hammock log file

2020-11-09 22:42:53.451: 65 sequences to be inserted into clusters
2020-11-09 22:42:53.451: 33 clusters to be extended
2020-11-09 22:42:53.475: 40 sequences rejected
2020-11-09 22:42:53.476: Overlap threshold is 0. Running full cluster merging.
2020-11-09 22:42:53.484: Buiding hhs...
2020-11-09 22:42:53.503: HH clustering...
2020-11-09 22:42:58.432:
Ready. Clustering time : 6857
2020-11-09 22:42:58.432: Resulting clusers: 52
2020-11-09 22:42:58.432: Containing 316 unique sequences and 430 total sequences.
2020-11-09 22:42:58.433: Unique sequences not assigned: 1036, total sequences not assigned: 1054
2020-11-09 22:42:58.434: Saving results to outupt files...
2020-11-09 22:42:58.462: Results in: /home/rstudio/data/HammockOrganoids/final_clusters_sequences.tsv
2020-11-09 22:42:58.462: and: /home/rstudio/data/HammockOrganoids/final_clusters.tsv
2020-11-09 22:42:58.463: and: /home/rstudio/data/HammockOrganoids/final_clusters_sequences_original_order.tsv
2020-11-09 22:42:58.463:
Calculating KLD...
2020-11-09 22:42:58.552: Final system KLD over match state MSA positions: 14.648972765490813
2020-11-09 22:42:58.552: Final system KLD over all MSA positions: 16.32643297890421
2020-11-09 22:42:58.552: Program successfully ended.

Generation of Weblogo visualization

```
ham.clusters <- data.table(read.table("/home/rstudio/data/HammockOrganoids/final_clusters.tsv",
  header = TRUE, skip = 0, sep = "\t", stringsAsFactors = FALSE, fill = TRUE))
id.order <- as.list(ham.clusters$cluster_id)
ham.clusters.all <- data.table(read.table("/home/rstudio/data/HammockOrganoids/final_clusters_sequences.tsv",
  header = TRUE, skip = 0, sep = "\t", stringsAsFactors = FALSE, fill = TRUE))
ham.clusters.all[, `:=`(alignment, gsub("\\-", "\\_", alignment))]
setkey(select.samples, Peptide)
setkey(select.samples.trsp, Peptide)

unlink("/home/rstudio/data/WEBlogosOrganoids", recursive = TRUE, force = FALSE)
dir.create(file.path("/home/rstudio/data/", "WEBlogosOrganoids"), showWarnings = FALSE)
dir.create(file.path("/home/rstudio/data/HammockOrganoids/", "alignments_final_Scored"),
  showWarnings = FALSE)

setkey(ham.clusters.all, cluster_id)
setkey(ham.clusters, cluster_id)
setkey(select.samples.pepMerge, Peptide)

opts_chunk$set(out.width = "100%", fig.align = "center")
generateWeblogo <- function(in.name) {
  # in.name <- ham.clusters$cluster_id[2] in.name <- 4352
  print(in.name)
  this.fa <- read.fasta(file = paste("/home/rstudio/data/HammockOrganoids/alignments_final/",
    in.name, ".aln", sep = ""))
  allSeqs <- unlist(getSequence(this.fa, as.string = TRUE))
  allSeqs <- data.table(unlist(lapply(allSeqs, function(x) gsub("([ -])", "",
    toupper(x)))))
  allSeqs.out <- select.samples.pepMerge[J(allSeqs)]
  allSeqs.out$Annot <- data.table(getName(this.fa))
  allSeqs.out[, `:=`(Annot, paste(Annot, "_", Score, sep = ""))]
```

```

allSeqs.out$Alignment <- data.table(toupper(unlist(getSequence(this.fa,
  as.string = TRUE))))

allSeqs.out <- allSeqs.out[rep(1:.N, Score)][, `:=`(Indx, 1:.N), by = Peptide]
allSeqs.out[, `:=`(Annot, paste(Annot, "_", Indx, sep = ""))]

write.fasta(as.list(allSeqs.out$Alignment), allSeqs.out$Annot, nbchar = 60,
  paste("/home/rstudio/data/HamcockOrganoids/alignments_final_Scored/",
    in.name, ".aln", sep = ""), open = "w")

this.main <- ham.clusters[J(in.name)]
main.gene <- select.samples.trsp[J(this.main$main_sequence)]$GeneName[1]
this.title <- paste("## Peptide", this.main$main_sequence, "from", main.gene,
  "with cluster number", in.name, sep = " ")

tmp <- system(paste("weblogo --format PDF --sequence-type protein --size large --errorbars NO --resolution",
  this.title, "' < /home/rstudio/data/HamcockOrganoids/alignments_final_Scored/",
  in.name, ".aln > /home/rstudio/data/WEBlogosOrganoids/", in.name, ".pdf",
  sep = ""), intern = TRUE, ignore.stdout = FALSE)

cat("\n")
cat(this.title, "\n")
cat("\n")
cat("\n")
cat(paste0("![Peptide: ", this.main$main_sequence, " from ", main.gene,
  " with cluster number ", in.name, "](/home/rstudio/data/WEBlogosOrganoids/",
  in.name, ".pdf)"))
cat("\n")
out.table <- knitr::kable(this.main, format = "latex")
print(column_spec(out.table, 1:ncol(this.main), monospace = TRUE) %>% kable_styling(latex_options = c("striped",
  "scale_down", "repeat_header"))))
cat("\n")
this.cluster <- ham.clusters.all[J(in.name)]
out.table <- knitr::kable(this.cluster, format = "latex")
print(column_spec(out.table, 1:ncol(this.cluster), monospace = TRUE) %>%
  kable_styling(latex_options = c("striped", "scale_down", "repeat_header"))))

this.found <- select.samples[J(this.cluster$sequence)]
setnames(this.cluster, "sequence", "Peptide")
this.found <- merge(this.found, this.cluster[, 2:3], by = "Peptide", all = FALSE)
cat("\n")
cat("\n")
output.order <- c("alignment", "LUTnrs", "GeneName", "start", "structure",
  "Group", "Score")
if (nrow(this.found) >= 48) {
  this.found.p1 <- this.found[1:47, ]
  out.table <- knitr::kable(this.found.p1[, ..output.order], format = "latex")
  print(column_spec(out.table, 1, monospace = TRUE) %>% kable_styling(latex_options = c("striped",
    "scale_down", "repeat_header"))))
  cat("\n")
  cat("\n\n\\pagebreak\n")
  cat("\n\n\\clearpage\n")
  this.found <- this.found[48:nrow(this.found), ]
  if (nrow(this.found) >= 48) {
    this.found.p2 <- this.found[1:47, ]
    out.table <- knitr::kable(this.found.p2[, ..output.order], format = "latex")
    print(column_spec(out.table, 1, monospace = TRUE) %>% kable_styling(latex_options = c("striped",

```

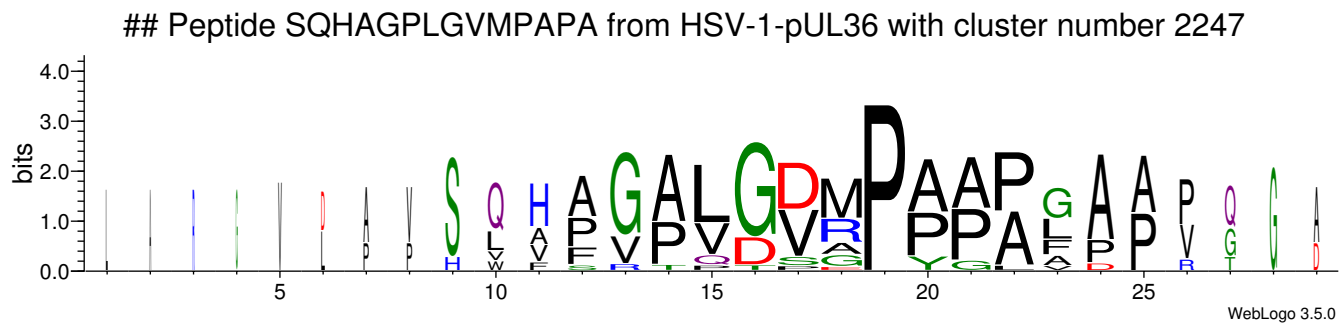
```

        "scale_down", "repeat_header"))))
cat("\n")
cat("\n\n\\pagebreak\n")
cat("\n\n\\clearpage\n")
this.found <- this.found[48:nrow(this.found), ]
}
if (nrow(this.found) >= 48) {
  this.found.p2 <- this.found[1:47, ]
  out.table <- knitr::kable(this.found.p2[, ..output.order], format = "latex")
  print(column_spec(out.table, 1, monospace = TRUE) %>% kable_styling(latex_options = c("striped",
    "scale_down", "repeat_header"))))
  cat("\n")
  cat("\n\n\\pagebreak\n")
  cat("\n\n\\clearpage\n")
  this.found <- this.found[48:nrow(this.found), ]
}
out.table <- knitr::kable(this.found[, ..output.order], format = "latex")
print(column_spec(out.table, 1, monospace = TRUE) %>% kable_styling(latex_options = c("striped",
  "scale_down", "repeat_header"))))
cat("\n")
cat("\n\n\\clearpage\n")
cat("\n")
} else {
  out.table <- knitr::kable(this.found[, ..output.order], format = "latex")
  print(column_spec(out.table, 1, monospace = TRUE) %>% kable_styling(latex_options = c("striped",
    "scale_down", "repeat_header"))))
  cat("\n")
  cat("\n\n\\pagebreak\n")
  cat("\n\n\\clearpage\n")
}
}
invisible(lapply(id.order, generateWeblogo))

```

[1] 2247

Peptide SQHAGPLGVMPAPA from HSV-1-pUL36 with cluster number 2247



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2247	SQHAGPLGVMPAPA	23	8	15

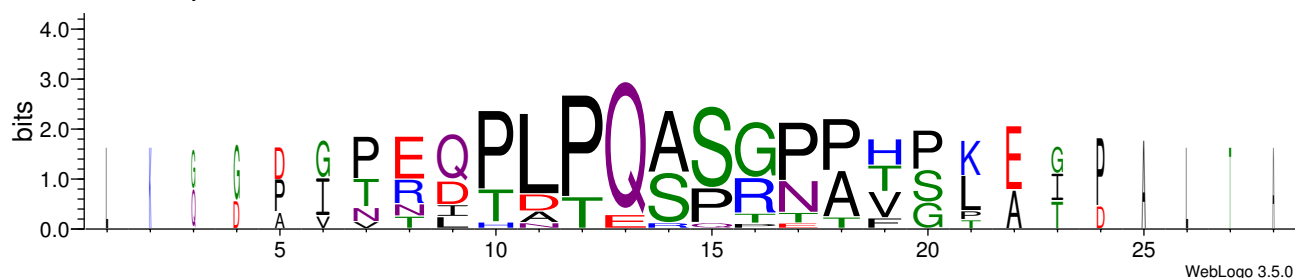
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2247	SQHAGPLGVMPAPA	-----SQHAGPLGVMPAPA-----	3	1	2
2247	SLVPGAQDSAPPGP	-----SLVPGAQDSAPPGP-----	2	1	1
2247	GVMPAPAGAAPQGA	-----GVMPAPAGAAPQGA-----	2	1	1
2247	AVGDRPPAPLAPVG	-----AVGDRPPAPLAPVG-----	2	1	1
2247	AGPLGVMPAPAGAA	-----AGPLGVMPAPAGAA-----	2	1	1
2247	VLPPhWAPGALDDGPYAPFPPR	___VLPPhWAPGALDDGPYAPFPPR___	1	1	0
2247	VFFVAVGDRPPAPA	-----VFFVAVGDRPPAPA-----	1	0	1
2247	VAVGDRPPAPLAPV	-----VAVGDRPPAPLAPV-----	1	0	1
2247	TPAPPALGAAPTGD	-----TPAPPALGAAPTGD-----	1	1	0
2247	SRTPGDEPAPAVAA	-----SRTPGDEPAPAVAA-----	1	0	1
2247	PGALDDGPYAPFPP	-----PGALDDGPYAPFPP-----	1	0	1
2247	LGVPAPAGAAPQG	-----LGVPAPAGAAPQG-----	1	0	1
2247	LARGVDAVSQHAGPLGVMPAPA	LARGVDAVSQHAGPLGVMPAPA-----	1	0	1
2247	FVAVGDRPPAPLAP	-----FVAVGDRPPAPLAP-----	1	0	1
2247	FVAVGDRPPAPADP	-----FVAVGDRPPAPADP-----	1	0	1
2247	AVSQHAGPLGVMPA	-----AVSQHAGPLGVMPA-----	1	1	0
2247	APGALDDGPYAPFP	-----APGALDDGPYAPFP-----	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
-----AGPLGVMPAPAGAA-----	X44054	HSV-1-pUL36	744	14aa	mRNA_30cpc_Organoid_MD114	1
-----AGPLGVMPAPAGAA-----	X44054	HSV-1-pUL36	744	14aa	mRNA_3000cpc_Organoid_MD101	1
-----APGALDDGPYAPFP-----	X41944	HSV-2-pUS4	301	14aa	mRNA_3000cpc_Organoid_MD101	1
-----AVGDRPPAPLAPVG-----	X44646	PRV-Becker-gE	205	14aa	mRNA_30cpc_Organoid_MD114	1
-----AVGDRPPAPLAPVG-----	X44646	PRV-TJ-gE	206	14aa	mRNA_30cpc_Organoid_MD114	1
-----AVGDRPPAPLAPVG-----	X44646	PRV-Becker-gE	205	14aa	mRNA_3000cpc_Organoid_MD101	1
-----AVGDRPPAPLAPVG-----	X44646	PRV-TJ-gE	206	14aa	mRNA_3000cpc_Organoid_MD101	1
-----AVSQHAGPLGVMPA-----	X56752	HSV-1-pUL36	739	14aaA5	mRNA_30cpc_Organoid_MD114	1
-----FVAVGDRPPAPADP-----	X76192	PRV-Becker-gE	203	14aa	mRNA_3000cpc_Organoid_MD101	1
-----FVAVGDRPPAPADP-----	X76192	PRV-TJ-gE	204	14aa	mRNA_3000cpc_Organoid_MD101	1
-----FVAVGDRPPAPLAP-----	X76191	PRV-Becker-gE	203	14aa	mRNA_3000cpc_Organoid_MD101	1
-----FVAVGDRPPAPLAP-----	X76191	PRV-TJ-gE	204	14aa	mRNA_3000cpc_Organoid_MD101	1
-----GVMPAPAGAAPQGA-----	X57955	HSV-1-pUL36	748	14aaA5	mRNA_30cpc_Organoid_MD114	1
-----GVMPAPAGAAPQGA-----	X57955	HSV-1-pUL36	748	14aaA5	mRNA_3000cpc_Organoid_MD101	1
LARGVDAVSQHAGPLGVMPAPA-----	X31969	HSV-1-pUL36	733	22aa	mRNA_3000cpc_Organoid_MD101	1
-----LGVPAPAGAAPQG-----	X32660	HSV-1-pUL36	747	14aa	mRNA_3000cpc_Organoid_MD101	1
-----PGALDDGPYAPFPP-----	X23987	HSV-2-pUS4	302	14aa	mRNA_3000cpc_Organoid_MD101	1
-----SLVPGAQDSAPPGP-----	X12339	HSV-1-pUL21	269	14aa	mRNA_30cpc_Organoid_MD114	1
-----SLVPGAQDSAPPGP-----	X12339	HSV-1-pUL21	269	14aa	mRNA_3000cpc_Organoid_MD101	1
-----SQHAGPLGVMPAPA-----	X11365	HSV-1-pUL36	741	14aa	mRNA_30cpc_Organoid_MD114	1
-----SQHAGPLGVMPAPA-----	X11365	HSV-1-pUL36	741	14aa	mRNA_3000cpc_Organoid_MD101	2
-----SRTPGDEPAPAVAA-----	X11807	BV-G	301	14aa	mRNA_3000cpc_Organoid_MD101	1
-----TPAPPALGAAPTGD-----	X7374	HSV-1-pUL27	53	14aa	mRNA_30cpc_Organoid_MD114	1
-----VAVGDRPPAPLAPV-----	X68849	PRV-Becker-gE	204	14aa	mRNA_3000cpc_Organoid_MD101	1
-----VAVGDRPPAPLAPV-----	X68849	PRV-TJ-gE	205	14aa	mRNA_3000cpc_Organoid_MD101	1
-----VFFVAVGDRPPAPA-----	X59346	PRV-Becker-gE	201	14aaA5	mRNA_3000cpc_Organoid_MD101	1
-----VFFVAVGDRPPAPA-----	X59346	PRV-TJ-gE	202	14aaA5	mRNA_3000cpc_Organoid_MD101	1
___VLPPhWAPGALDDGPYAPFPPR___	X67733	HSV-2-pUS4	295	22aa	mRNA_30cpc_Organoid_MD114	1

[1] 2339

Peptide PEQPLPQASRNPTG from HIV-tat-GR17 with cluster number 2339

Peptide PEQPLPQASRNPTG from HIV-tat-GR17 with cluster number 2339



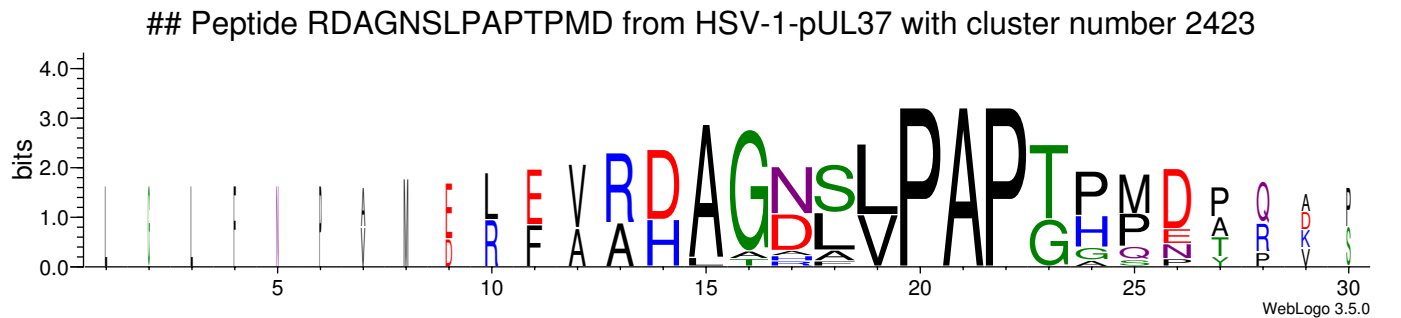
cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2339	PEQPLPQASRNPTG	21	9	12

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2339	PEQPLPQASRNPTG	____PEQPLPQASRNPTG_____	3	1	2
2339	TTIPAPQASTTFFP	____TTIPAPQASTTFFP_____	1	0	1
2339	TLPQSPGPAVPLAG	____TLPQSPGPAVPLAG_____	1	0	1
2339	TIPAPQASTTFFPT	____TIPAPQASTTFFPT_____	1	1	0
2339	TEASGPTHKEIPA	____TEASGPTHKEIPA_____	1	0	1
2339	RDITLQSPGPAVPL	____RDITLQSPGPAVPL_____	1	1	0
2339	QPLPQASRNPTGPE	____QPLPQASRNPTGPE_____	1	0	1
2339	QDPIPEQPLPQASR	____QDPIPEQPLPQASR_____	1	0	1
2339	PIPEQPLPQASRNP	____PIPEQPLPQASRNP_____	1	0	1
2339	PDTQSSGPAHSKET	____PDTQSSGPAHSKET_____	1	1	0
2339	NTEASGPTHKEIP	____NTEASGPTHKEIP_____	1	1	0
2339	NNLPDTQSSGPAHSKETPALTA	____NNLPDTQSSGPAHSKETPALTA_____	1	0	1
2339	NNLPDTQSSGPAHS	____NNLPDTQSSGPAHS_____	1	1	0
2339	LPQSPGPAVPLAGD	____LPQSPGPAVPLAGD_____	1	1	0
2339	LKGGDGTDTLPQSPGPAVPLA	____LKGGDGTDTLPQSPGPAVPLA_____	1	0	1
2339	IPEQPLPQASRNPT	____IPEQPLPQASRNPT_____	1	1	0
2339	GTRDTLPQSPGPAV	____GTRDTLPQSPGPAV_____	1	0	1
2339	GDGTRDTLPQSPGP	____GDGTRDTLPQSPGP_____	1	0	1
2339	AVVEQHLPQRQPEP	____AVVEQHLPQRQPEP_____	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
____AVVEQHLPQRQPEP_____	X56833	HSV-2-pUS8	379	14aaA5	mRNA_30cpc_Organoid_MD114	1
____GDGTRDTLPQSPGP_____	X57646	HSV-1-pUS4	97	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____GTRDTLPQSPGPAV_____	X61567	HSV-1-pUS4	99	14aa	mRNA_3000cpc_Organoid_MD101	1
____IPEQPLPQASRNPT_____	X15454	HIV-tat-GR17	69	14aa	mRNA_30cpc_Organoid_MD114	1
____LKGGDGTDTLPQSPGPAVPLA_____	X28727	HSV-1-pUS4	94	22aa	mRNA_3000cpc_Organoid_MD101	1
____LPQSPGPAVPLAGD_____	X30196	HSV-1-pUS4	104	14aa	mRNA_30cpc_Organoid_MD114	1
____NNLPDTQSSGPAHS_____	X142	PV2-VP1	25	14aa	mRNA_30cpc_Organoid_MD114	1
____NNLPDTQSSGPAHSKETPALTA_____	X141	PV2-VP1	25	22aa	mRNA_3000cpc_Organoid_MD101	1
____NTEASGPTHKEIP_____	X528	PV1-VP1	29	14aa	mRNA_30cpc_Organoid_MD114	1
____PDTQSSGPAHSKET_____	X22894	PV2-VP1	28	14aa	mRNA_30cpc_Organoid_MD114	1
____PEQPLPQASRNPTG_____	X23115	HIV-tat-GR17	70	14aa	mRNA_30cpc_Organoid_MD114	1
____PEQPLPQASRNPTG_____	X23115	HIV-tat-GR17	70	14aa	mRNA_3000cpc_Organoid_MD101	1
____PEQPLPQASRNPTG_____	X51225	HIV-tat-GR17	70	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____PIPEQPLPQASRNP_____	X21435	HIV-tat-GR17	68	14aa	mRNA_3000cpc_Organoid_MD101	1
____QDPIPEQPLPQASR_____	X19677	HIV-tat-GR17	66	14aa	mRNA_3000cpc_Organoid_MD101	1
____QPLPQASRNPTGPE_____	X19276	HIV-tat-GR17	72	14aa	mRNA_3000cpc_Organoid_MD101	1
____RDTLPQSPGPAVPL_____	X26647	HSV-1-pUS4	101	14aa	mRNA_30cpc_Organoid_MD114	1
____TEASGPTHKEIPA_____	X8328	PV1-VP1	30	14aa	mRNA_3000cpc_Organoid_MD101	1
____TIPAPQASTTFFPT_____	X6623	HSV-1-pUS7	220	14aa	mRNA_30cpc_Organoid_MD114	1
____TLPQSPGPAVPLAG_____	X7811	HSV-1-pUS4	103	14aa	mRNA_3000cpc_Organoid_MD101	1
____TTIPAPQASTTFFP_____	X5986	HSV-1-pUS7	219	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2423

Peptide RDAGNSLPAPTPMD from HSV-1-pUL37 with cluster number 2423



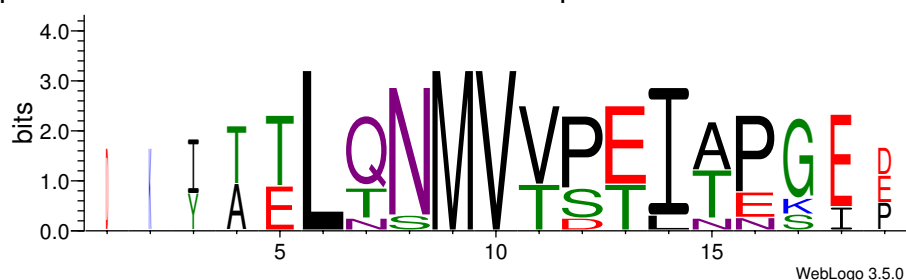
cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2423	RDAGNSLPAPTPMD	19	8	11

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2423	RDAGNSLPAPTPMD	-----RDAGNSLPAPTPMD-----	3	1	2
2423	LEVRDAGNSLPAPT	-----LEVRDAGNSLPAPT-----	2	1	1
2423	VRDAGNSLPAPTPM	-----VRDAGNSLPAPTPM-----	1	1	0
2423	VMERFAAHAGDLVPAPGHPDPR	-----VMERFAAHAGDLVPAPGHPDPR-----	1	1	0
2423	RFAAHAGDLVPAPG	-----RFAAHAGDLVPAPG-----	1	1	0
2423	RALPAPTGGNTQVS	-----RALPAPTGGNTQVS-----	1	1	0
2423	LTHALPAPTGGNTQ	-----LTHALPAPTGGNTQ-----	1	0	1
2423	LGLFNPAMERFAAHAGDLVPAP	LGLFNPAMERFAAHAGDLVPAP-----	1	1	0
2423	HAGDLVPAPGHPEP	-----HAGDLVPAPGHPEP-----	1	0	1
2423	GDLVPAPGHPDPRA	-----GDLVPAPGHPDPRA-----	1	0	1
2423	EVRDAGNSLPAPTP	-----EVRDAGNSLPAPTP-----	1	0	1
2423	DLEVRDAGNSLPAPTPMDAQKP	-----DLEVRDAGNSLPAPTPMDAQKP-----	1	0	1
2423	DAGNSLPAPTPMDA	-----DAGNSLPAPTPMDA-----	1	0	1
2423	AHAGDLVPAPGHPE	-----AHAGDLVPAPGHPE-----	1	1	0
2423	AHAGDLVPAPGHPD	-----AHAGDLVPAPGHPD-----	1	0	1
2423	AAFVPAPGASPYPD	-----AAFVPAPGASPYPD-----	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
_____AAFVPAPGASPYPD_	X56629	HSV-1-pUL36	943	14aaA5	mRNA_3000cpc_Organoid_MD101	1
_____AHAGDLVPAPGHPD_	X41345	HSV-1-pUL19	418	14aa	mRNA_3000cpc_Organoid_MD101	1
_____AHAGDLVPAPGHPD_	X41345	HSV-2-pUL19	418	14aa	mRNA_3000cpc_Organoid_MD101	1
_____AHAGDLVPAPGHPE_	X55989	HSV-1-pUL19	418	14aaA5	mRNA_30cpc_Organoid_MD114	1
_____AHAGDLVPAPGHPE_	X55989	HSV-2-pUL19	418	14aaA5	mRNA_30cpc_Organoid_MD114	1
_____DAGNSLPAPTPMDA_	X35985	HSV-1-pUL37	1062	14aa	mRNA_3000cpc_Organoid_MD101	1
_____DLEVRDAGNSLPAPTPMDAQKP	X35377	HSV-1-pUL37	1057	22aa	mRNA_3000cpc_Organoid_MD101	1
_____EVRDAGNSLPAPTP_	X39316	HSV-1-pUL37	1059	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GDLVPAPGHPDPRA_	X63627	HSV-1-pUL19	421	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GDLVPAPGHPDPRA_	X63627	HSV-2-pUL19	421	14aa	mRNA_3000cpc_Organoid_MD101	1
_____HAGDLVPAPGHPEP_	X18063	HSV-1-pUL19	419	14aa	mRNA_3000cpc_Organoid_MD101	1
_____HAGDLVPAPGHPEP_	X18063	HSV-2-pUL19	419	14aa	mRNA_3000cpc_Organoid_MD101	1
_____LEVRDAGNSLPAPT_	X31743	HSV-1-pUL37	1058	14aa	mRNA_30cpc_Organoid_MD114	1
_____LEVRDAGNSLPAPT_	X31743	HSV-1-pUL37	1058	14aa	mRNA_3000cpc_Organoid_MD101	1
LGLFNPAMERFAAHAGDLVPAP_	X32520	HSV-1-pUL19	406	22aa	mRNA_30cpc_Organoid_MD114	1
LGLFNPAMERFAAHAGDLVPAP_	X32520	HSV-2-pUL19	406	22aa	mRNA_30cpc_Organoid_MD114	1
_____LTHALPAPTGGNTQ_	X28893	EV71-VP1-1095-Japan-97	20	14aa	mRNA_3000cpc_Organoid_MD101	1
_____LTHALPAPTGGNTQ_	X28893	EV71-VP1-FY23	20	14aa	mRNA_3000cpc_Organoid_MD101	1
_____LTHALPAPTGGNTQ_	X28893	EV71-VP1-cx2011-016	20	14aa	mRNA_3000cpc_Organoid_MD101	1
_____RALPAPTGGNTQVS_	X27174	EV71-VP1-1095-Japan-97	22	14aa	mRNA_30cpc_Organoid_MD114	1
_____RALPAPTGGNTQVS_	X27174	EV71-VP1-FY23	22	14aa	mRNA_30cpc_Organoid_MD114	1
_____RALPAPTGGNTQVS_	X27174	EV71-VP1-cx2011-016	22	14aa	mRNA_30cpc_Organoid_MD114	1
_____RDAGNSLPAPTPMD_	X26767	HSV-1-pUL37	1061	14aa	mRNA_30cpc_Organoid_MD114	1
_____RDAGNSLPAPTPMD_	X26767	HSV-1-pUL37	1061	14aa	mRNA_3000cpc_Organoid_MD101	2
_____RFAAHAGDLVPAPG_	X28278	HSV-1-pUL19	415	14aa	mRNA_30cpc_Organoid_MD114	1
_____RFAAHAGDLVPAPG_	X28278	HSV-2-pUL19	415	14aa	mRNA_30cpc_Organoid_MD114	1
_____VMERFAAHAGDLVPAPGHDPDR_	X66824	HSV-1-pUL19	412	22aa	mRNA_30cpc_Organoid_MD114	1
_____VMERFAAHAGDLVPAPGHDPDR_	X66824	HSV-2-pUL19	412	22aa	mRNA_30cpc_Organoid_MD114	1
_____VRDAGNSLPAPTPM_	X67519	HSV-1-pUL37	1060	14aa	mRNA_30cpc_Organoid_MD114	1

Peptide TLQNMVVPEIAPGE from HSV-1-pUL19 with cluster number 2140

ptide TLQNMVVPEIAPGE from HSV-1-pUL19 with cluster number



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2140	TLQNMVVPEIAPGE	19	10	9

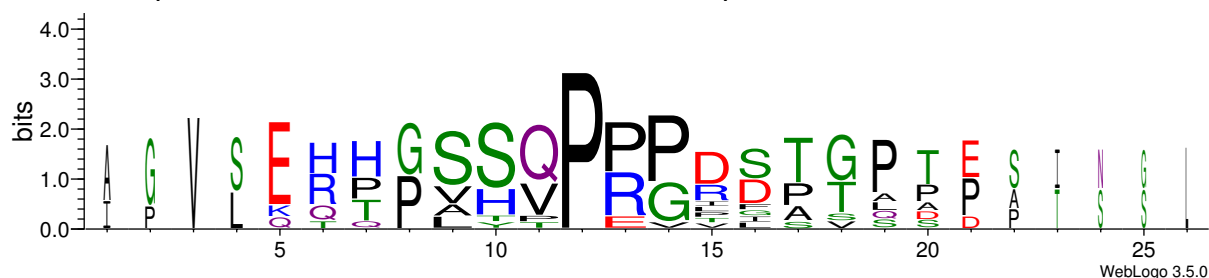
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2140	TLQNMVVPEIAPGE	____TLQNMVVPEIAPGE_	4	2	2
2140	YATLQNMVVPEITP	_YATLQNMVVPEITP__	2	0	2
2140	TELTNMVTSTITEK	___TELTNMVTSTITEK__	2	1	1
2140	LQNMVVPEITPGED	____LQNMVVPEITPGED	2	1	1
2140	LQNMVVPEIAPGEE	____LQNMVVPEIAPGEE	2	2	0
2140	LNSMVTDTLNSIP	____LNSMVTDTLNSIP	2	1	1
2140	ITELTNMVTSTITE	_ITELTNMVTSTITE__	2	2	0
2140	ATLQNMVVPEIAPG	___ATLQNMVVPEIAPG__	2	1	1
2140	DKITELTNMVTSTI	DKITELTNMVTSTI_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
___ATLQNMVVPEIAPG__	X40328	HSV-1-pUL19	837	14aa	mRNA_30cpc_Organoid_MD114	1
___ATLQNMVVPEIAPG__	X40328	HSV-2-pUL19	837	14aa	mRNA_30cpc_Organoid_MD114	1
___ATLQNMVVPEIAPG__	X40328	HSV-1-pUL19	837	14aa	mRNA_3000cpc_Organoid_MD101	1
___ATLQNMVVPEIAPG__	X40328	HSV-2-pUL19	837	14aa	mRNA_3000cpc_Organoid_MD101	1
DKITELTNMVTSTI_____	X33879	PV1-VP1	474	14aa	mRNA_3000cpc_Organoid_MD101	1
_ITELTNMVTSTITE__	X14645	PV1-VP1	476	14aa	mRNA_30cpc_Organoid_MD114	2
____LNSMVTDTLNSIP	X28344	BoNT-E-Hc	31	14aa	mRNA_30cpc_Organoid_MD114	1
____LNSMVTDTLNSIP	X28344	BoNT-E-Hc	31	14aa	mRNA_3000cpc_Organoid_MD101	1
____LQNMVVPEIAPGEE	X29938	HSV-1-pUL19	839	14aa	mRNA_30cpc_Organoid_MD114	2
____LQNMVVPEIAPGEE	X29938	HSV-2-pUL19	839	14aa	mRNA_30cpc_Organoid_MD114	2
____LQNMVVPEITPGED	X29937	HSV-1-pUL19	839	14aa	mRNA_30cpc_Organoid_MD114	1
____LQNMVVPEITPGED	X29937	HSV-2-pUL19	839	14aa	mRNA_30cpc_Organoid_MD114	1
____LQNMVVPEITPGED	X29937	HSV-1-pUL19	839	14aa	mRNA_3000cpc_Organoid_MD101	1
____LQNMVVPEITPGED	X29937	HSV-2-pUL19	839	14aa	mRNA_3000cpc_Organoid_MD101	1
___TELTNMVTSTITEK__	X8305	PV1-VP1	477	14aa	mRNA_30cpc_Organoid_MD114	1
___TELTNMVTSTITEK__	X8305	PV1-VP1	477	14aa	mRNA_3000cpc_Organoid_MD101	1
____TLQNMVVPEIAPGE_	X7795	HSV-1-pUL19	838	14aa	mRNA_30cpc_Organoid_MD114	1
____TLQNMVVPEIAPGE_	X47264	HSV-1-pUL19	838	14aaA5	mRNA_30cpc_Organoid_MD114	1
____TLQNMVVPEIAPGE_	X7795	HSV-2-pUL19	838	14aa	mRNA_30cpc_Organoid_MD114	1
____TLQNMVVPEIAPGE_	X47264	HSV-2-pUL19	838	14aaA5	mRNA_30cpc_Organoid_MD114	1
____TLQNMVVPEIAPGE_	X7795	HSV-1-pUL19	838	14aa	mRNA_3000cpc_Organoid_MD101	1
____TLQNMVVPEIAPGE_	X47264	HSV-1-pUL19	838	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____TLQNMVVPEIAPGE_	X7795	HSV-2-pUL19	838	14aa	mRNA_3000cpc_Organoid_MD101	1
____TLQNMVVPEIAPGE_	X47264	HSV-2-pUL19	838	14aaA5	mRNA_3000cpc_Organoid_MD101	1
_YATLQNMVVPEITP__	X71450	HSV-1-pUL19	836	14aa	mRNA_3000cpc_Organoid_MD101	2
_YATLQNMVVPEITP__	X71450	HSV-2-pUL19	836	14aa	mRNA_3000cpc_Organoid_MD101	2

[1] 2739

Peptide AGVSERTPVSVPPP from HSV-2-pUS8 with cluster number 2739

Peptide AGVSERTPVSVPPP from HSV-2-pUS8 with cluster number 2739



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2739	AGVSERTPVSVPPP	17	9	8

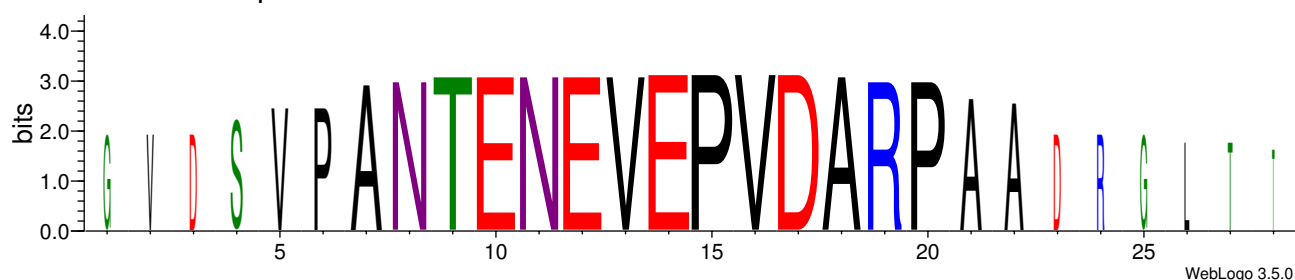
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2739	SSQPRGDSTGPTE	_____SSQPRGDSTGPTE_____	2	1	1
2739	EHHGSHQPPRPDPT	____EHHGSHQPPRPDPT_____	2	2	0
2739	AGVSERTPVSVPPP	AGVSERTPVSVPPP_____	2	1	1
2739	VLEHHGSHQPPRPD	__VLEHHGSHQPPRPD_____	1	0	1
2739	TQASPPEGPGTGPS	____TQASPPEGPGTGPS_____	1	0	1
2739	QRHGLYVPRDDPVLADAINGL	____QRHGLYVPRDDPVLADAINGL_____	1	0	1
2739	QPSSQPRGDSTGPT	____QPSSQPRGDSTGPT_____	1	0	1
2739	PSSQPRGDSTGPTE	____PSSQPRGDSTGPTE_____	1	1	0
2739	PRVLFSTQPPPTSS	____PRVLFSTQPPPTSS_____	1	1	0
2739	PLTTPPPIIATSDP	____PLTTPPPIIATSDP_____	1	1	0
2739	PASVPEPVLASAPP	____PASVPEPVLASAPP_____	1	0	1
2739	KQPSSQPRGDSTGP	____KQPSSQPRGDSTGP_____	1	1	0
2739	IPVLEHHGSHQPPP	____IPVLEHHGSHQPPP_____	1	1	0
2739	GVSERTPVSVPPPT	____GVSERTPVSVPPPT_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
AGVSERTPVSVPPP_____	X44254	HSV-2-pUS8	178	14aa	mRNA_30cpc_Organoid_MD114	1
AGVSERTPVSVPPP_____	X44254	HSV-2-pUS8	178	14aa	mRNA_3000cpc_Organoid_MD101	1
____EHHGSHQPPRPDPT_____	X54998	HSV-1-pUL44	455	14aaA5	mRNA_30cpc_Organoid_MD114	1
____EHHGSHQPPRPDPT_____	X86288	HSV-1-pUL44	455	14aaG4S	mRNA_30cpc_Organoid_MD114	1
____EHHGSHQPPRPDPT_____	X54998	HSV-2-pUL44	424	14aaA5	mRNA_30cpc_Organoid_MD114	1
____EHHGSHQPPRPDPT_____	X86288	HSV-2-pUL44	424	14aaG4S	mRNA_30cpc_Organoid_MD114	1
_GVSERTPVSVPPPT_____	X64797	HSV-2-pUS8	179	14aa	mRNA_3000cpc_Organoid_MD101	1
IPVLEHHGSHQPPP_____	X80594	HSV-1-pUL44	451	14aaG4S	mRNA_30cpc_Organoid_MD114	1
IPVLEHHGSHQPPP_____	X80594	HSV-2-pUL44	420	14aaG4S	mRNA_30cpc_Organoid_MD114	1
____KQPSSQPRGDSTGP_____	X3874	HIV-tat	71	14aa	mRNA_30cpc_Organoid_MD114	1
____PASVPEPVLASAPP_____	X23374	HSV-1-pUL36	445	14aa	mRNA_3000cpc_Organoid_MD101	1
____PLTTPPPIIATSDP_____	X22562	HSV-1-pUL1	191	14aa	mRNA_30cpc_Organoid_MD114	1
____PRVLFSTQPPPTSS_____	X22500	NGF	67	14aa	mRNA_30cpc_Organoid_MD114	1
____PSSQPRGDSTGPTE_____	X21128	HIV-tat	73	14aa	mRNA_30cpc_Organoid_MD114	1
____QPSSQPRGDSTGPT_____	X19191	HIV-tat	72	14aa	mRNA_3000cpc_Organoid_MD101	1
____QRHGLYVPRDDPVLADAINGL_____	X19358	HSV-1-pUL21	376	22aa	mRNA_3000cpc_Organoid_MD101	1
____SSQPRGDSTGPTE_____	X10707	HIV-tat	74	14aa	mRNA_30cpc_Organoid_MD114	1
____SSQPRGDSTGPTE_____	X10707	HIV-tat	74	14aa	mRNA_3000cpc_Organoid_MD101	1
____TQASPPEGPGTGPS_____	X7067	Tau	275	14aa	mRNA_3000cpc_Organoid_MD101	1
____VLEHHGSHQPPRPD_____	X67856	HSV-1-pUL44	453	14aa	mRNA_3000cpc_Organoid_MD101	1
____VLEHHGSHQPPRPD_____	X67856	HSV-2-pUL44	422	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2503

Peptide ANTENEVEPVDARP from APP with cluster number 2503

Peptide ANTENEVEPVDARP from APP with cluster number 2503



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2503	ANTENEVEPVDARP	17	4	13

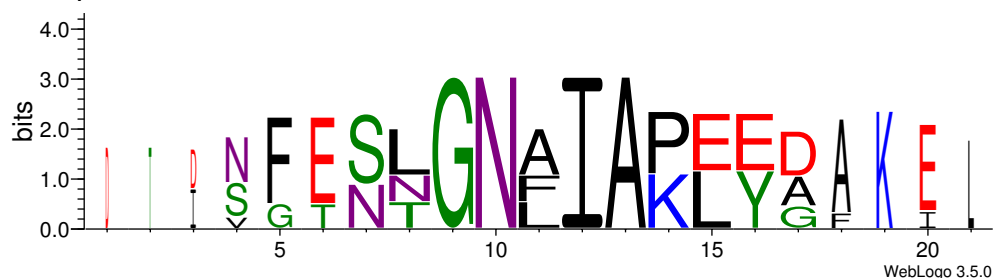
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2503	ANTENEVEPVDARP	_____ANTENEVEPVDARP_____	5	3	2
2503	GVDSVPANTENEVEPVDARPA	GVDSVPANTENEVEPVDARPA_____	3	0	3
2503	VPANTENEVEPVDA	____VPANTENEVEPVDA_____	2	0	2
2503	TENEVEPVDARPA	____TENEVEPVDARPA_____	2	1	1
2503	SVPANTENEVEPVDARPAADRG	___SVPANTENEVEPVDARPAADRG___	1	0	1
2503	SVPANTENEVEPVD	___SVPANTENEVEPVD_____	1	0	1
2503	NTENEVEPVDARPA	____NTENEVEPVDARPA_____	1	0	1
2503	EPVDARPAADRGLT	____EPVDARPAADRGLT_____	1	0	1
2503	ANTENEVEPVDARPAADRGLTT	____ANTENEVEPVDARPAADRGLTT	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
_____ANTENEVEPVDARP_____	X39803	APP	631	14aa	mRNA_30cpc_Organoid_MD114	1
_____ANTENEVEPVDARP_____	X55577	APP	631	14aaA5	mRNA_30cpc_Organoid_MD114	1
_____ANTENEVEPVDARP_____	X86867	APP	631	14aaG4S	mRNA_30cpc_Organoid_MD114	1
_____ANTENEVEPVDARP_____	X39803	APP	631	14aa	mRNA_3000cpc_Organoid_MD101	2
_____ANTENEVEPVDARPAADRGLTT	X39804	APP	631	22aa	mRNA_3000cpc_Organoid_MD101	1
_____EPVDARPAADRGLT_____	X37901	APP	638	14aa	mRNA_3000cpc_Organoid_MD101	1
GVDSVPANTENEVEPVDARPA_____	X64911	APP	625	22aa	mRNA_3000cpc_Organoid_MD101	3
_____NTENEVEPVDARPA_____	X524	APP	632	14aa	mRNA_3000cpc_Organoid_MD101	1
___SVPANTENEVEPVD_____	X48790	APP	628	14aaA5	mRNA_3000cpc_Organoid_MD101	1
___SVPANTENEVEPVDARPAADRG___	X13535	APP	628	22aa	mRNA_3000cpc_Organoid_MD101	1
_____TENEVEPVDARPA_____	X8197	APP	633	14aa	mRNA_30cpc_Organoid_MD114	1
_____TENEVEPVDARPA_____	X8197	APP	633	14aa	mRNA_3000cpc_Organoid_MD101	1
_____VPANTENEVEPVDA_____	X67260	APP	629	14aa	mRNA_3000cpc_Organoid_MD101	2

[1] 2794

Peptide NFESNGNFIAP EYA from H5N1-HA with cluster number 2794

Peptide NFESNGNFIAP EYA from H5N1-HA with cluster number 2794



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2794	NFESNGNFIAP EYA	15	6	9

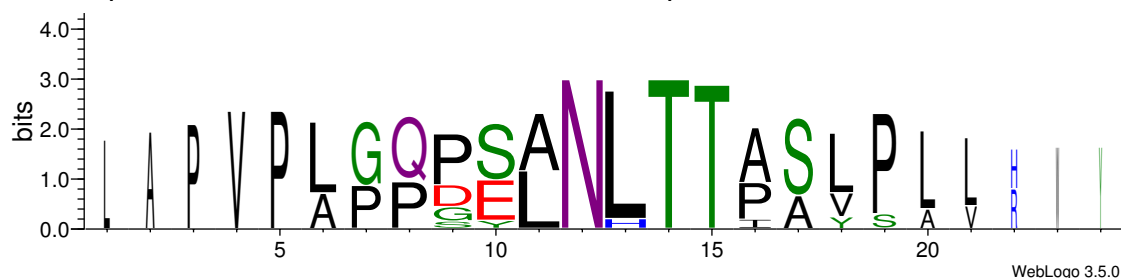
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2794	NFESNGNFIAP EYA	__NFESNGNFIAP EYA__	4	1	3
2794	NLGNAIAKLEDAKE	____NLGNAIAKLEDAKE_	3	1	2
2794	SFESTGNLIAPEYG	___SFESTGNLIAPEYG___	2	1	1
2794	LGNAIAKLEDAKEL	____LGNAIAKLEDAKEL	2	1	1
2794	STGNLIAPEYGFKI	____STGNLIAPEYGFKI_	1	1	0
2794	GTNLGNAIAKLEDA	____GTNLGNAIAKLEDA__	1	0	1
2794	DVGTLNLAIAKLE	__DVGTLNLAIAKLE____	1	1	0
2794	DTISFESTGNLIAP	DTISFESTGNLIAP_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
DTISFESTGNLIAP_____	X34028	H2N2-HA	251	14aa	mRNA_3000cpc_Organoid_MD101	1
__DVGTLNLAIAKLE____	X36365	MV-F-CAM-70	461	14aa	mRNA_30cpc_Organoid_MD114	1
__DVGTLNLAIAKLE____	X36365	MV-F-Edmonston	461	14aa	mRNA_30cpc_Organoid_MD114	1
__DVGTLNLAIAKLE____	X36365	MV-F-Halle	461	14aa	mRNA_30cpc_Organoid_MD114	1
____GTNLGNAIAKLEDA__	X57096	MV-F-CAM-70	463	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____GTNLGNAIAKLEDA__	X57096	MV-F-Edmonston	463	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____GTNLGNAIAKLEDA__	X57096	MV-F-Halle	463	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____LGNAIAKLEDAKEL	X32295	MV-F-CAM-70	466	14aa	mRNA_30cpc_Organoid_MD114	1
____LGNAIAKLEDAKEL	X32295	MV-F-Edmonston	466	14aa	mRNA_30cpc_Organoid_MD114	1
____LGNAIAKLEDAKEL	X32295	MV-F-Halle	466	14aa	mRNA_30cpc_Organoid_MD114	1
____LGNAIAKLEDAKEL	X32295	MV-F-CAM-70	466	14aa	mRNA_3000cpc_Organoid_MD101	1
____LGNAIAKLEDAKEL	X32295	MV-F-Edmonston	466	14aa	mRNA_3000cpc_Organoid_MD101	1
____LGNAIAKLEDAKEL	X32295	MV-F-Halle	466	14aa	mRNA_3000cpc_Organoid_MD101	1
___NFESNGNFIAP EYA___	X2967	H5N1-HA	256	14aa	mRNA_30cpc_Organoid_MD114	1
___NFESNGNFIAP EYA___	X2967	H5N1-HA	256	14aa	mRNA_3000cpc_Organoid_MD101	2
___NFESNGNFIAP EYA___	X46009	H5N1-HA	256	14aaA5	mRNA_3000cpc_Organoid_MD101	1
____NLGNAIAKLEDAKE_	X1704	MV-F-CAM-70	465	14aa	mRNA_30cpc_Organoid_MD114	1
____NLGNAIAKLEDAKE_	X1704	MV-F-Edmonston	465	14aa	mRNA_30cpc_Organoid_MD114	1
____NLGNAIAKLEDAKE_	X1704	MV-F-Halle	465	14aa	mRNA_30cpc_Organoid_MD114	1
____NLGNAIAKLEDAKE_	X1704	MV-F-CAM-70	465	14aa	mRNA_3000cpc_Organoid_MD101	2
____NLGNAIAKLEDAKE_	X1704	MV-F-Edmonston	465	14aa	mRNA_3000cpc_Organoid_MD101	2
____NLGNAIAKLEDAKE_	X1704	MV-F-Halle	465	14aa	mRNA_3000cpc_Organoid_MD101	2
___SFESTGNLIAPEYG___	X14087	H2N2-HA	254	14aa	mRNA_30cpc_Organoid_MD114	1
___SFESTGNLIAPEYG___	X14087	H2N2-HA	254	14aa	mRNA_3000cpc_Organoid_MD101	1
____STGNLIAPEYGFKI_	X10468	H2N2-HA	257	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2277

Peptide APPGELNLTTASVP from HSV-1-pUL22 with cluster number 2277

Peptide APPGELNLTTASVP from HSV-1-pUL22 with cluster number 2277



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2277	APPGELNLTTASVP	14	4	10

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2277	VPLGQPSANLTPA	__VPLGQPSANLTPA_____	2	0	2
2277	LAPVPLGQPSANLT	LAPVPLGQPSANLT_____	2	0	2
2277	APPGELNLTTASVP	____APPGELNLTTASVP____	2	1	1
2277	SYLNHTTISLPAVH	_______SYLNHTTISLPAVH__	1	1	0
2277	QPSANLTPAYSLL	_______QPSANLTPAYSLL____	1	1	0
2277	PVPLGQPSANLTP	__PVPLGQPSANLTP_____	1	0	1
2277	PPDELNLTTASLPL	_______PPDELNLTTASLPL____	1	0	1
2277	PDELNLTTASLPLL	_______PDELNLTTASLPLL____	1	0	1
2277	LNLTTASLPLLRWY	_______LNLTTASLPLLRWY____	1	1	0
2277	APVPLGQPSANLTT	_APVPLGQPSANLTT_____	1	0	1
2277	APPDELNLTTASLP	_______APPDELNLTTASLP____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
____APPDELNLTTASLP____	X41723	HSV-1-pUL22	67	14aa	mRNA_3000cpc_Organoid_MD101	1
____APPDELNLTTASLP____	X41723	HSV-2-pUL22	67	14aa	mRNA_3000cpc_Organoid_MD101	1
____APPGELNLTTASVP____	X56090	HSV-1-pUL22	67	14aaA5	mRNA_30cpc_Organoid_MD114	1
____APPGELNLTTASVP____	X56090	HSV-2-pUL22	67	14aaA5	mRNA_30cpc_Organoid_MD114	1
____APPGELNLTTASVP____	X41744	HSV-1-pUL22	67	14aa	mRNA_3000cpc_Organoid_MD101	1
____APPGELNLTTASVP____	X41744	HSV-2-pUL22	67	14aa	mRNA_3000cpc_Organoid_MD101	1
_APVPLGQPSANLTT_____	X41962	HSV-1-pUL37	206	14aa	mRNA_3000cpc_Organoid_MD101	1
LAPVPLGQPSANLT_____	X31942	HSV-1-pUL37	205	14aa	mRNA_3000cpc_Organoid_MD101	1
LAPVPLGQPSANLT_____	X53552	HSV-1-pUL37	205	14aaA5	mRNA_3000cpc_Organoid_MD101	1
_______LNLTTASLPLLRWY____	X28419	HSV-1-pUL22	72	14aa	mRNA_30cpc_Organoid_MD114	1
_______LNLTTASLPLLRWY____	X28419	HSV-2-pUL22	72	14aa	mRNA_30cpc_Organoid_MD114	1
_______PDELNLTTASLPLL____	X22994	HSV-1-pUL22	69	14aa	mRNA_3000cpc_Organoid_MD101	1
_______PDELNLTTASLPLL____	X22994	HSV-2-pUL22	69	14aa	mRNA_3000cpc_Organoid_MD101	1
_______PPDELNLTTASLPL____	X22100	HSV-1-pUL22	68	14aa	mRNA_3000cpc_Organoid_MD101	1
_______PPDELNLTTASLPL____	X22100	HSV-2-pUL22	68	14aa	mRNA_3000cpc_Organoid_MD101	1
__PVPLGQPSANLTP_____	X24172	HSV-1-pUL37	207	14aa	mRNA_3000cpc_Organoid_MD101	1
_______QPSANLTPAYSLL____	X19197	HSV-1-pUL37	212	14aa	mRNA_30cpc_Organoid_MD114	1
_______SYLNHTTISLPAVH__	X13793	BDV-CRNP5	60	14aa	mRNA_30cpc_Organoid_MD114	1
_______SYLNHTTISLPAVH__	X13793	BDV-G	60	14aa	mRNA_30cpc_Organoid_MD114	1
_______SYLNHTTISLPAVH__	X13793	BDV-No-98	60	14aa	mRNA_30cpc_Organoid_MD114	1
____VPLGQPSANLTPA_____	X67210	HSV-1-pUL37	208	14aa	mRNA_3000cpc_Organoid_MD101	2

[1] 2255

Peptide AHAVSNAINAGSAA from HSV-1-pUL36 with cluster number 2255

Peptide AHAVSNAINAGSAA from HSV-1-pUL36 with cluster number 2255



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2255	AHAVSNAINAGSAA	12	6	6

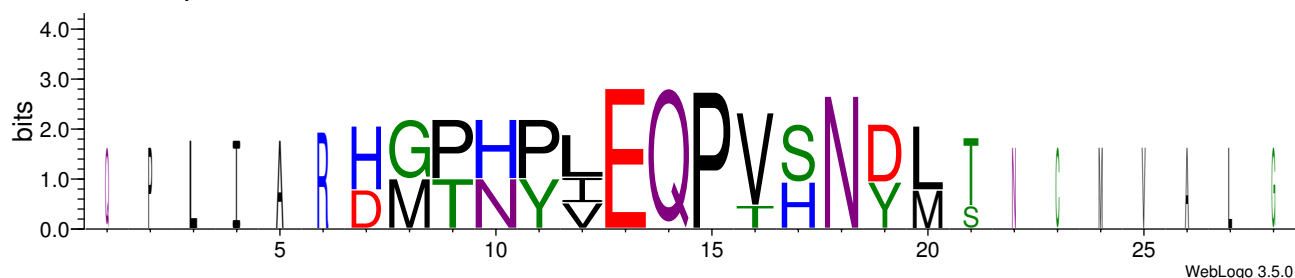
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2255	AHAVSNAINAGSAA	___AHAVSNAINAGSAA_____	4	2	2
2255	VSNAINAGSAAVAT	_____VSNAINAGSAAVAT_____	3	1	2
2255	AHAVSNAINAGSAAVATADPAS	___AHAVSNAINAGSAAVATADPAS___	2	1	1
2255	VSNAINAGSAAVATADPASTVD	_____VSNAINAGSAAVATADPASTVD	1	1	0
2255	PIIAHAVSNAINAGSAAVATAD	PIIAHAVSNAINAGSAAVATAD_____	1	0	1
2255	INAGSAAVATADPA	_____INAGSAAVATADPA_____	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
___AHAVSNAINAGSAA_____	X41348	HSV-1-pUL36	1138	14aa	mRNA_30cpc_Organoid_MD114	2
___AHAVSNAINAGSAA_____	X41348	HSV-1-pUL36	1138	14aa	mRNA_3000cpc_Organoid_MD101	2
___AHAVSNAINAGSAAVATADPAS___	X41349	HSV-1-pUL36	1138	22aa	mRNA_30cpc_Organoid_MD114	1
___AHAVSNAINAGSAAVATADPAS___	X41349	HSV-1-pUL36	1138	22aa	mRNA_3000cpc_Organoid_MD101	1
_____INAGSAAVATADPA_____	X14290	HSV-1-pUL36	1145	14aa	mRNA_30cpc_Organoid_MD114	1
PIIAHAVSNAINAGSAAVATAD_____	X21423	HSV-1-pUL36	1135	22aa	mRNA_3000cpc_Organoid_MD101	1
_____VSNAINAGSAAVAT_____	X66274	HSV-1-pUL36	1141	14aa	mRNA_30cpc_Organoid_MD114	1
_____VSNAINAGSAAVAT_____	X66274	HSV-1-pUL36	1141	14aa	mRNA_3000cpc_Organoid_MD101	2
_____VSNAINAGSAAVATADPASTVD	X66275	HSV-1-pUL36	1141	22aa	mRNA_30cpc_Organoid_MD114	1

[1] 2831

Peptide HMTNYLEQPVSNDL from MV-H-CAM-70 with cluster number 2831

Peptide HMTNYLEQPVSNDL from MV-H-CAM-70 with cluster number 2831



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2831	HMTNYLEQPVSNDL	11	7	4

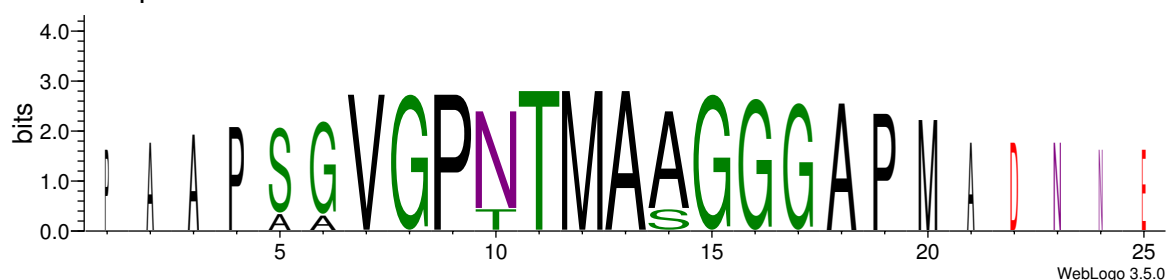
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2831	HMTNYLEQPVSNDL	-----HMTNYLEQPVSNDL-----	4	2	2
2831	GPHPVEQPVHNYMT	-----GPHPVEQPVHNYMT-----	2	1	1
2831	RDGPHPVEQPVHNY	-----RDGPHPVEQPVHNY-----	1	1	0
2831	QPLIARDGPHPIEQ	QPLIARDGPHPIEQ-----	1	1	0
2831	LIARDGPHPIEQPT	__LIARDGPHPIEQPT-----	1	0	1
2831	HMTNYLEQPVSNDLSNCMVALG	-----HMTNYLEQPVSNDLSNCMVALG	1	1	0
2831	GPHPIEQPTHNYMT	-----GPHPIEQPTHNYMT-----	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
-----GPHPIEQPTHNYMT-----	X62586	HSV-1-pUL19	108	14aa	mRNA_30cpc_Organoid_MD114	1
-----GPHPVEQPVHNYMT-----	X62589	HSV-2-pUL19	108	14aa	mRNA_30cpc_Organoid_MD114	1
-----GPHPVEQPVHNYMT-----	X62589	HSV-2-pUL19	108	14aa	mRNA_3000cpc_Organoid_MD101	1
-----HMTNYLEQPVSNDL-----	X17313	MV-H-CAM-70	271	14aa	mRNA_30cpc_Organoid_MD114	2
-----HMTNYLEQPVSNDL-----	X17313	MV-H-Edmonston	271	14aa	mRNA_30cpc_Organoid_MD114	2
-----HMTNYLEQPVSNDL-----	X17313	MV-H-Halle	271	14aa	mRNA_30cpc_Organoid_MD114	2
-----HMTNYLEQPVSNDL-----	X17313	MV-H-CAM-70	271	14aa	mRNA_3000cpc_Organoid_MD101	2
-----HMTNYLEQPVSNDL-----	X17313	MV-H-Edmonston	271	14aa	mRNA_3000cpc_Organoid_MD101	2
-----HMTNYLEQPVSNDL-----	X17313	MV-H-Halle	271	14aa	mRNA_3000cpc_Organoid_MD101	2
-----HMTNYLEQPVSNDLSNCMVALG	X17312	MV-H-CAM-70	271	22aa	mRNA_30cpc_Organoid_MD114	1
-----HMTNYLEQPVSNDLSNCMVALG	X17312	MV-H-Edmonston	271	22aa	mRNA_30cpc_Organoid_MD114	1
-----HMTNYLEQPVSNDLSNCMVALG	X17312	MV-H-Halle	271	22aa	mRNA_30cpc_Organoid_MD114	1
-----HMTNYLEQPVSNDLSNCMVALG	X17312	MV-H-MVi-Illinois.USA-50.99	271	22aa	mRNA_30cpc_Organoid_MD114	1
-----HMTNYLEQPVSNDLSNCMVALG	X17312	MV-H-MVi-Texas.USA-28.99	271	22aa	mRNA_30cpc_Organoid_MD114	1
__LIARDGPHPIEQPT-----	X29647	HSV-1-pUL19	103	14aa	mRNA_3000cpc_Organoid_MD101	1
QPLIARDGPHPIEQ-----	X19271	HSV-1-pUL19	101	14aa	mRNA_30cpc_Organoid_MD114	1
QPLIARDGPHPIEQ-----	X19271	HSV-2-pUL19	101	14aa	mRNA_30cpc_Organoid_MD114	1
-----RDGPHPVEQPVHNY-----	X26778	HSV-2-pUL19	106	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2785

Peptide VGPNTMAAGGGAPM from AAV8-VP1 with cluster number 2785

Peptide VGPNTMAAGGGAPM from AAV8-VP1 with cluster number 2785



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2785	VGPNTMAAGGGAPM	11	6	5

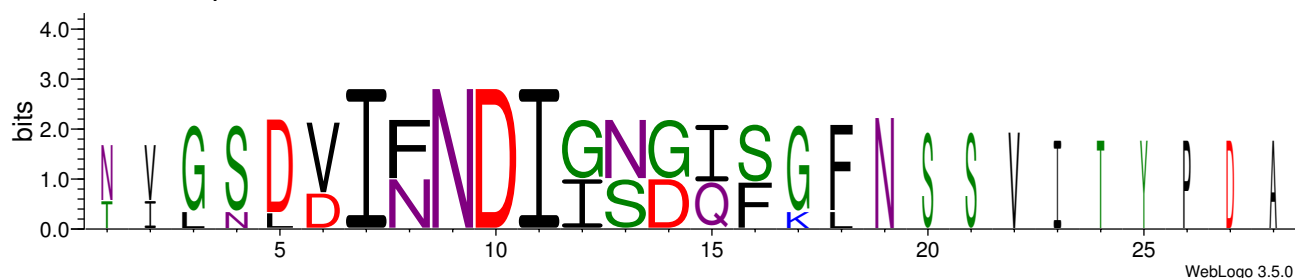
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2785	VGPNTMAAGGGAPM	-----VGPNTMAAGGGAPM-----	3	2	1
2785	TTMASGGGAPMADN	-----TTMASGGGAPMADN--	1	0	1
2785	SGVGPNTMAAGGGA	---SGVGPNTMAAGGGA-----	1	0	1
2785	PSGVPNTMAAGGGAPMADNNE	___PSGVPNTMAAGGGAPMADNNE	1	1	0
2785	PAAPSGVGPNTMAA	PAAPSGVGPNTMAA-----	1	0	1
2785	GVGPNTMAAGGGAP	-----GVGPNTMAAGGGAP-----	1	1	0
2785	APSGVGPNTMAAGG	--APSGVGPNTMAAGG-----	1	0	1
2785	AAVGPTTMASGGGA	---AAVGPTTMASGGGA-----	1	1	0
2785	AAPSGVGPNTMAAG	_AAPSGVGPNTMAAG-----	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
_AAPSGVGPNTMAAG-----	X43545	AAV8-VP1	194	14aa	mRNA_30cpc_Organoid_MD114	1
---AAVGPTTMASGGGA-----	X87907	AAV1-VP1	196	14aaG4S	mRNA_30cpc_Organoid_MD114	1
---AAVGPTTMASGGGA-----	X87907	AAV6-VP1	196	14aaG4S	mRNA_30cpc_Organoid_MD114	1
--APSGVGPNTMAAGG-----	X41614	AAV8-VP1	195	14aa	mRNA_3000cpc_Organoid_MD101	1
-----GVGPNTMAAGGGAP-----	X64974	AAV8-VP1	198	14aa	mRNA_30cpc_Organoid_MD114	1
PAAPSGVGPNTMAA-----	X23600	AAV8-VP1	193	14aa	mRNA_3000cpc_Organoid_MD101	1
___PSGVPNTMAAGGGAPMADNNE	X21320	AAV1-VP1	195	22aa	mRNA_30cpc_Organoid_MD114	1
___PSGVPNTMAAGGGAPMADNNE	X21320	AAV6-VP1	195	22aa	mRNA_30cpc_Organoid_MD114	1
___PSGVPNTMAAGGGAPMADNNE	X21320	AAV8-VP1	196	22aa	mRNA_30cpc_Organoid_MD114	1
---SGVGPNTMAAGGGA-----	X13360	AAV8-VP1	197	14aa	mRNA_3000cpc_Organoid_MD101	1
-----TTMASGGGAPMADN--	X5996	AAV1-VP1	201	14aa	mRNA_3000cpc_Organoid_MD101	1
-----TTMASGGGAPMADN--	X5996	AAV6-VP1	201	14aa	mRNA_3000cpc_Organoid_MD101	1
-----TTMASGGGAPMADN--	X5996	AAV8-VP1	202	14aa	mRNA_3000cpc_Organoid_MD101	1
-----TTMASGGGAPMADN--	X5996	AAV9-VP1	202	14aa	mRNA_3000cpc_Organoid_MD101	1
-----VGPNTMAAGGGAPM-----	X69003	AAV8-VP1	199	14aa	mRNA_30cpc_Organoid_MD114	2
-----VGPNTMAAGGGAPM-----	X69003	AAV8-VP1	199	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2597

Peptide GSDVIFNDIGNGQF from BoNT-G-Hc with cluster number 2597

Peptide GSDVIFNDIGNGQF from BoNT-G-Hc with cluster number 2597



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2597	GSDVIFNDIGNGQF	11	5	6

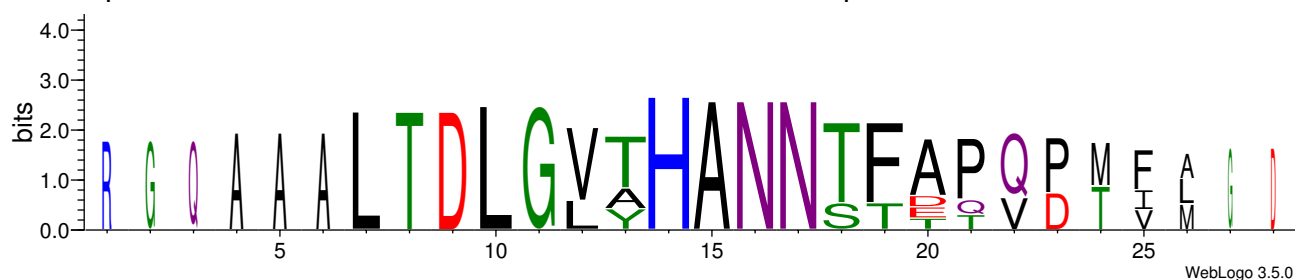
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2597	GSDVIFNDIGNGQF	__GSDVIFNDIGNGQF_____	3	1	2
2597	NVGSDVIFNDIGNG	NVGSDVIFNDIGNG_____	2	1	1
2597	INNDIISDISGFNSSVITYPDA	_____INNDIISDISGFNSSVITYPDA	2	1	1
2597	DINNDIISDISGFN	_____DINNDIISDISGFN_____	2	2	0
2597	TILNLDINNDIISDISGFNSSV	TILNLDINNDIISDISGFNSSV_____	1	0	1
2597	DVIFNDIGNGQFKL	_____DVIFNDIGNGQFKL_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
_____DINNDIISDISGFN_____	X34449	TeNT-Hc	27	14aa	mRNA_30cpc_Organoid_MD114	2
_____DVIFNDIGNGQFKL_____	X36259	BoNT-G-Hc	29	14aa	mRNA_3000cpc_Organoid_MD101	1
__GSDVIFNDIGNGQF_____	X61902	BoNT-G-Hc	27	14aa	mRNA_30cpc_Organoid_MD114	1
__GSDVIFNDIGNGQF_____	X61902	BoNT-G-Hc	27	14aa	mRNA_3000cpc_Organoid_MD101	2
_____INNDIISDISGFNSSVITYPDA	X14169	TeNT-Hc	28	22aa	mRNA_30cpc_Organoid_MD114	1
_____INNDIISDISGFNSSVITYPDA	X14169	TeNT-Hc	28	22aa	mRNA_3000cpc_Organoid_MD101	1
NVGSDVIFNDIGNG_____	X45919	BoNT-G-Hc	25	14aaA5	mRNA_30cpc_Organoid_MD114	1
NVGSDVIFNDIGNG_____	X45919	BoNT-G-Hc	25	14aaA5	mRNA_3000cpc_Organoid_MD101	1
TILNLDINNDIISDISGFNSSV_____	X6634	TeNT-Hc	22	22aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2012

Peptide AAALTDLGVTHANNTFAPQPMF from HSV-1-pUL35 with cluster number 2012

Peptide AAALTDLGVTHANNTFAPQPMF from HSV-1-pUL35 with cluster number 2012



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2012	AAALTDLGVTHANNTFAPQPMF	9	4	5

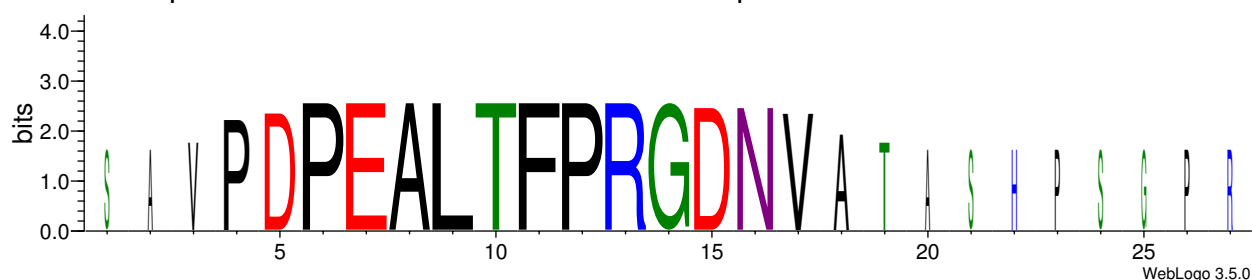
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2012	YHANNSTEQVDTIM	-----YHANNSTEQVDTIM__	1	1	0
2012	YHANNSTDVDTVL	-----YHANNSTDVDTVL__	1	0	1
2012	RGQAAALTDLGVTHANNTFAPQ	RGQAAALTDLGVTHANNTFAPQ_____	1	0	1
2012	RGQAAALTDLGLAH	RGQAAALTDLGLAH_____	1	1	0
2012	LTDLGVTHANNTFAPQPMFAGD	-----LTDLGVTHANNTFAPQPMFAGD	1	0	1
2012	LTDLGVTHANNTFA	-----LTDLGVTHANNTFA_____	1	0	1
2012	LTDLGLAHANNTFT	-----LTDLGLAHANNTFT_____	1	0	1
2012	LGVTHANNTFAPQP	-----LGVTHANNTFAPQP_____	1	1	0
2012	AAALTDLGVTHANNTFAPQPMF	___AAALTDLGVTHANNTFAPQPMF___	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
___AAALTDLGVTHANNTFAPQPMF___	X43752	HSV-1-pUL35	58	22aa	mRNA_30cpc_Organoid_MD114	1
___AAALTDLGVTHANNTFAPQPMF___	X43752	HSV-2-pUL34	58	22aa	mRNA_30cpc_Organoid_MD114	1
-----LGVTHANNTFAPQP-----	X32656	HSV-2-pUL34	64	14aa	mRNA_30cpc_Organoid_MD114	1
-----LTDLGLAHANNTFT-----	X28986	HSV-1-pUL35	61	14aa	mRNA_3000cpc_Organoid_MD101	1
-----LTDLGVTHANNTFA-----	X52773	HSV-2-pUL34	61	14aaA5	mRNA_3000cpc_Organoid_MD101	1
-----LTDLGVTHANNTFAPQPMFAGD	X28987	HSV-1-pUL35	61	22aa	mRNA_3000cpc_Organoid_MD101	1
-----LTDLGVTHANNTFAPQPMFAGD	X28987	HSV-2-pUL34	61	22aa	mRNA_3000cpc_Organoid_MD101	1
RGQAAALTDLGLAH_____	X27444	HSV-1-pUL35	55	14aa	mRNA_30cpc_Organoid_MD114	1
RGQAAALTDLGLAH_____	X27444	HSV-2-pUL34	55	14aa	mRNA_30cpc_Organoid_MD114	1
RGQAAALTDLGVTHANNTFAPQ_____	X27447	HSV-1-pUL35	55	22aa	mRNA_3000cpc_Organoid_MD101	1
RGQAAALTDLGVTHANNTFAPQ_____	X27447	HSV-2-pUL34	55	22aa	mRNA_3000cpc_Organoid_MD101	1
-----YHANNSTDVDTVL__	X70760	H1N1-HA	24	14aa	mRNA_3000cpc_Organoid_MD101	1
-----YHANNSTEQVDTIM__	X70763	H2N2-HA	22	14aa	mRNA_30cpc_Organoid_MD114	1
-----YHANNSTEQVDTIM__	X70763	H5N1-HA	23	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2834

Peptide PDPEALTFPRGDNV from HSV-1-pUL22 with cluster number 2834

Peptide PDPEALTFPRGDNV from HSV-1-pUL22 with cluster number 2834



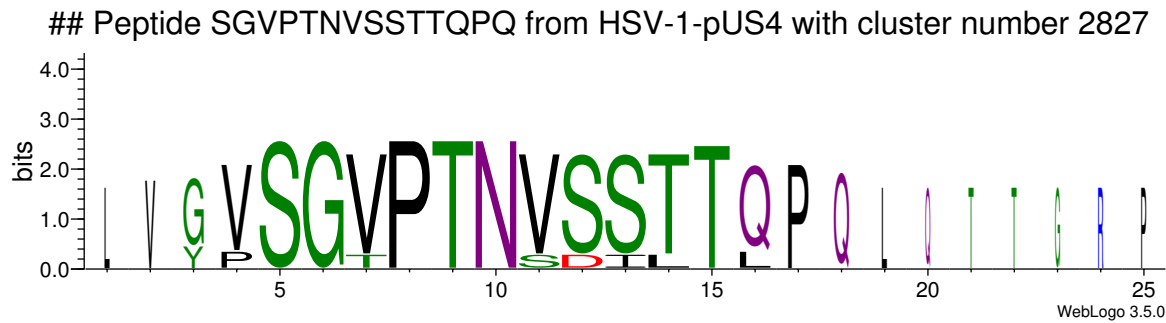
cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2834	PDPEALTFPRGDNV	8	5	3

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2834	PDPEALTFPRGDNV	___PDPEALTFPRGDNV_____	3	2	1
2834	VPDPEALTFPRGDN	__VPDPEALTFPRGDN_____	1	1	0
2834	SAVPDPEALTFPRG	SAVPDPEALTFPRG_____	1	0	1
2834	PEALTFPRGDNVATASHPSGPR	_____PEALTFPRGDNVATASHPSGPR	1	1	0
2834	PEALTFPRGDNVAT	_____PEALTFPRGDNVAT_____	1	0	1
2834	DPEALTFPRGDNVA	____DPEALTFPRGDNVA_____	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
____DPEALTFPRGDNVA_____	X34957	HSV-1-pUL22	168	14aa	mRNA_30cpc_Organoid_MD114	1
___PDPEALTFPRGDNV_____	X22953	HSV-1-pUL22	167	14aa	mRNA_30cpc_Organoid_MD114	2
___PDPEALTFPRGDNV_____	X22953	HSV-1-pUL22	167	14aa	mRNA_3000cpc_Organoid_MD101	1
_____PEALTFPRGDNVAT_____	X23217	HSV-1-pUL22	169	14aa	mRNA_3000cpc_Organoid_MD101	1
_____PEALTFPRGDNVATASHPSGPR	X23218	HSV-1-pUL22	169	22aa	mRNA_30cpc_Organoid_MD114	1
SAVPDPEALTFPRG_____	X13061	HSV-1-pUL22	164	14aa	mRNA_3000cpc_Organoid_MD101	1
__VPDPEALTFPRGDN_____	X67221	HSV-1-pUL22	166	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2827

Peptide SGVPTNVSSTTQPQ from HSV-1-pUS4 with cluster number 2827



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2827	SGVPTNVSSTTQPQ	8	3	5

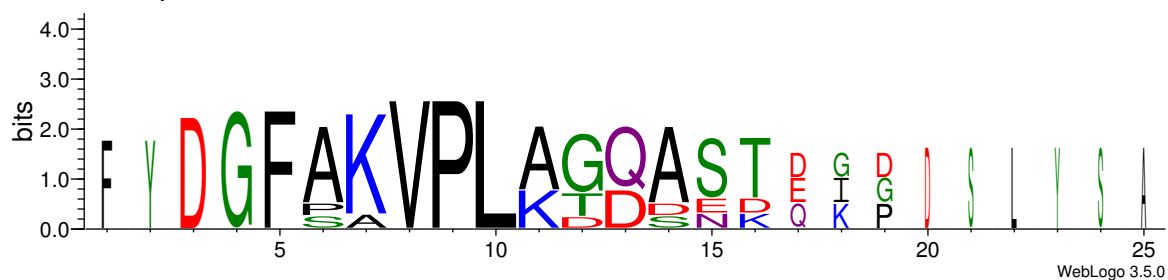
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2827	SGVPTNVSSTTQPQ	___SGVPTNVSSTTQPQ_____	2	2	0
2827	YPSGTPTNSDILTL	__YPSGTPTNSDILTL_____	1	0	1
2827	VSGVPTNVSSTTQPQLQTGRP	___VSGVPTNVSSTTQPQLQTGRP	1	0	1
2827	VSGVPTNVSSTTQP	___VSGVPTNVSSTTQP_____	1	0	1
2827	VGVS	__VGVS_____	1	0	1
2827	LVGV	__LVGV_____	1	1	0
2827	GV	__GV_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
__GVSGVPTNVSSTTQ_____	X64803	HSV-1-pUS4	21	14aa	mRNA_3000cpc_Organoid_MD101	1
LVGVSGVPTNVSST_____	X33077	HSV-1-pUS4	19	14aa	mRNA_30cpc_Organoid_MD114	1
___SGVPTNVSSTTQPQ_____	X13347	HSV-1-pUS4	23	14aa	mRNA_30cpc_Organoid_MD114	2
_VGVS_____	X69188	HSV-1-pUS4	20	14aa	mRNA_3000cpc_Organoid_MD101	1
___VSGVPTNVSSTTQP_____	X66506	HSV-1-pUS4	22	14aa	mRNA_3000cpc_Organoid_MD101	1
___VSGVPTNVSSTTQPQLQTGRP	X66505	HSV-1-pUS4	22	22aa	mRNA_3000cpc_Organoid_MD101	1
__YPSGTPTNSDILTL_____	X70871	TMEV-VP3	198	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2650

Peptide DGFAKVPLAGQAST from PV2-VP1 with cluster number 2650

Peptide DGFAKVPLAGQAST from PV2-VP1 with cluster number 2650



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2650	DGFAKVPLAGQAST	8	6	2

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2650	DGFAKVPLAGQAST	__DGFAKVPLAGQAST_____	3	2	1
2650	PAVPLAGDDEKDKP	____PAVPLAGDDEKDKP_____	1	1	0
2650	GFAKVPLKTDANDQIGDSLYSA	___GFAKVPLKTDANDQIGDSLYSA	1	1	0
2650	FYDGFSAKVLKDKS	FYDGFSAKVLKDKS_____	1	1	0
2650	FYDGFSAKVLKTD	FYDGFSAKVLKTD_____	1	1	0
2650	AKVPLAGQASTEGD	____AKVPLAGQASTEGD_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
____AKVPLAGQASTEGD_____	X40095	PV2-VP1	213	14aa	mRNA_3000cpc_Organoid_MD101	1
__DGFAKVPLAGQAST_____	X36211	PV2-VP1	210	14aa	mRNA_30cpc_Organoid_MD114	2
__DGFAKVPLAGQAST_____	X36211	PV2-VP1	210	14aa	mRNA_3000cpc_Organoid_MD101	1
FYDGFSAKVLKTD_____	X60984	PV3-VP1	784	14aaA5	mRNA_30cpc_Organoid_MD114	1
FYDGFSAKVLKDKS_____	X76306	PV1-VP1	208	14aa	mRNA_30cpc_Organoid_MD114	1
___GFAKVPLKTDANDQIGDSLYSA	X65528	PV3-VP1	787	22aa	mRNA_30cpc_Organoid_MD114	1
____PAVPLAGDDEKDKP_____	X23690	HSV-1-pUS4	110	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2830

Peptide AVGLLVLAGLVAAF from HSV-1-pUL27 with cluster number 2830

Peptide AVGLLVLAGLVAAF from HSV-1-pUL27 with cluster number 2830



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2830	AVGLLVLAGLVAAF	7	6	1

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2830	QTGVRVALALVAAFALAMAVLR	QTGVRVALALVAAFALAMAVLR__	1	1	0
2830	QTGVRVALALVAAF	QTGVRVALALVAAF_____	1	1	0
2830	NENDVLVALVASSP	_NENDVLVALVASSP_____	1	0	1
2830	ICALVVGALVA AVASAAPAAPA	_ICALVVGALVA AVASAAPAAPA_	1	1	0
2830	GLLVLAGLVAAFFAFRYVLQLQ	__GLLVLAGLVAAFFAFRYVLQLQ	1	1	0
2830	GLLVLAGLVAAFFA	__GLLVLAGLVAAFFA_____	1	1	0
2830	AVGLLVLAGLVAAF	AVGLLVLAGLVAAF_____	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
AVGLLVLAGLVAAF_____	X44643	HSV-1-pUL27	779	14aa	mRNA_30cpc_Organoid_MD114	1
AVGLLVLAGLVAAF_____	X44643	HSV-2-pUL27	776	14aa	mRNA_30cpc_Organoid_MD114	1
__GLLVLAGLVAAFFA_____	X57542	HSV-1-pUL27	781	14aaA5	mRNA_30cpc_Organoid_MD114	1
__GLLVLAGLVAAFFA_____	X57542	HSV-2-pUL27	778	14aaA5	mRNA_30cpc_Organoid_MD114	1
__GLLVLAGLVAAFFAFRYVLQLQ	X63279	HSV-1-pUL27	781	22aa	mRNA_30cpc_Organoid_MD114	1
__GLLVLAGLVAAFFAFRYVLQLQ	X63279	HSV-2-pUL27	778	22aa	mRNA_30cpc_Organoid_MD114	1
ICALVVGALVA AVASAAPAAPA	X16834	HSV-2-pUL27	7	22aa	mRNA_30cpc_Organoid_MD114	1
_NENDVLVALVASSP_____	X45739	HSV-1-pUL36	2098	14aaA5	mRNA_3000cpc_Organoid_MD101	1
QTGVRVALALVAAF_____	X50108	HSV-1-pUL10	319	14aaA5	mRNA_30cpc_Organoid_MD114	1
QTGVRVALALVAAF_____	X50108	HSV-2-pUL10	319	14aaA5	mRNA_30cpc_Organoid_MD114	1
QTGVRVALALVA AFALAMAVLR__	X18706	HSV-1-pUL10	319	22aa	mRNA_30cpc_Organoid_MD114	1
QTGVRVALALVA AFALAMAVLR__	X18706	HSV-2-pUL10	319	22aa	mRNA_30cpc_Organoid_MD114	1

[1] 2821

Peptide PATGDVHAMGALPG from AAV1-VP1 with cluster number 2821

Peptide PATGDVHAMGALPG from AAV1-VP1 with cluster number 2821



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2821	PATGDVHAMGALPG	7	1	6

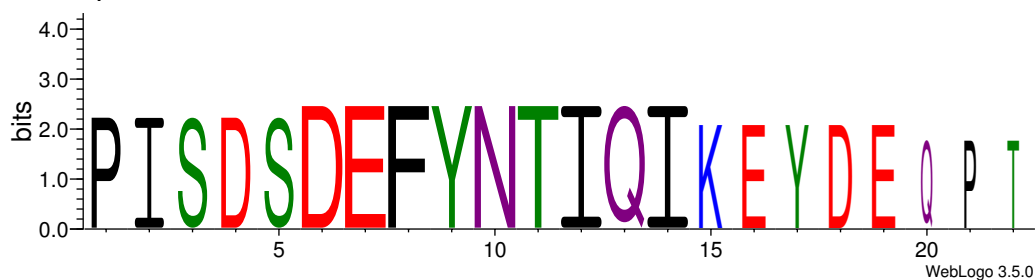
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2821	PATGDVHAMGALPG	____PATGDVHAMGALPG____	2	0	2
2821	TGDVHAMGALPGMV	____TGDVHAMGALPGMV____	1	0	1
2821	TDPATGDVHAMGAL	____TDPATGDVHAMGAL____	1	1	0
2821	SSSTDPATGDVHAMG	____SSSTDPATGDVHAMG____	1	0	1
2821	SSSTDPATGDVHVM	____SSSTDPATGDVHVM____	1	0	1
2821	SSSTDPATGDVHAMGALPGMVW	____SSSTDPATGDVHAMGALPGMVW____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
____PATGDVHAMGALPG____	X23344	AAV1-VP1	591	14aa	mRNA_3000cpc_Organoid_MD101	2
____PATGDVHAMGALPG____	X23344	AAV6-VP1	591	14aa	mRNA_3000cpc_Organoid_MD101	2
SSSTDPATGDVHAMGALPGMVW	X10627	AAV1-VP1	586	22aa	mRNA_3000cpc_Organoid_MD101	1
SSSTDPATGDVHAMGALPGMVW	X10627	AAV6-VP1	586	22aa	mRNA_3000cpc_Organoid_MD101	1
SSSTDPATGDVHVM_____	X10628	AAV1-VP1	586	14aa	mRNA_3000cpc_Organoid_MD101	1
SSSTDPATGDVHVM_____	X10628	AAV6-VP1	586	14aa	mRNA_3000cpc_Organoid_MD101	1
_SSSTDPATGDVHAMG_____	X10606	AAV1-VP1	587	14aa	mRNA_3000cpc_Organoid_MD101	1
_SSSTDPATGDVHAMG_____	X10606	AAV6-VP1	587	14aa	mRNA_3000cpc_Organoid_MD101	1
____TDPATGDVHAMGAL____	X78635	AAV1-VP1	589	14aaG4S	mRNA_30cpc_Organoid_MD114	1
____TDPATGDVHAMGAL____	X78635	AAV6-VP1	589	14aaG4S	mRNA_30cpc_Organoid_MD114	1
____TGDVHAMGALPGMV____	X8895	AAV1-VP1	593	14aa	mRNA_3000cpc_Organoid_MD101	1
____TGDVHAMGALPGMV____	X8895	AAV6-VP1	593	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2578

Peptide PISDSDEFYNTIQI from BoNT-B-Hc with cluster number 2578

Peptide PISDSDEFYNTIQI from BoNT-B-Hc with cluster number 2578



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2578	PISDSDEFYNTIQI	7	4	3

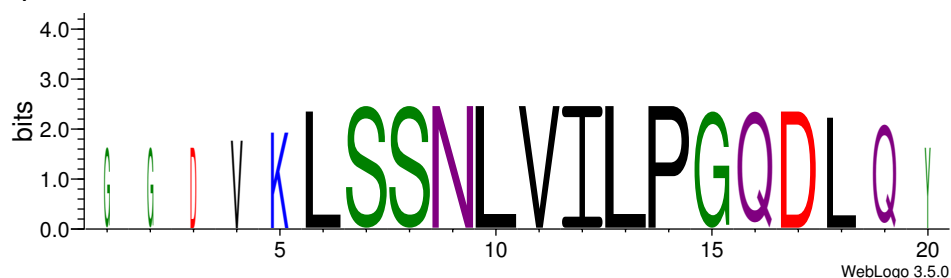
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2578	PISDSDEFYNTIQI	PISDSDEFYNTIQI_____	3	1	2
2578	PISDSDEFYNTIQIKEYDEQPT	PISDSDEFYNTIQIKEYDEQPT	2	1	1
2578	DEFYNTIQIKEYDE	____DEFYNTIQIKEYDE__	2	2	0

alignment	LUThrs	GeneName	start	structure	Group	Score
____DEFYNTIQIKEYDE__	X35767	BoNT-B-Hc	378	14aa	mRNA_30cpc_Organoid_MD114	2
PISDSDEFYNTIQI_____	X21416	BoNT-B-Hc	373	14aa	mRNA_30cpc_Organoid_MD114	1
PISDSDEFYNTIQI_____	X50793	BoNT-B-Hc	373	14aaA5	mRNA_3000cpc_Organoid_MD101	2
PISDSDEFYNTIQIKEYDEQPT	X21415	BoNT-B-Hc	373	22aa	mRNA_30cpc_Organoid_MD114	1
PISDSDEFYNTIQIKEYDEQPT	X21415	BoNT-B-Hc	373	22aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2544

Peptide LSSNLVILPGQDLQ from MV-H-CAM-70 with cluster number 2544

Peptide LSSNLVILPGQDLQ from MV-H-CAM-70 with cluster number 2



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2544	LSSNLVILPGQDLQ	7	2	5

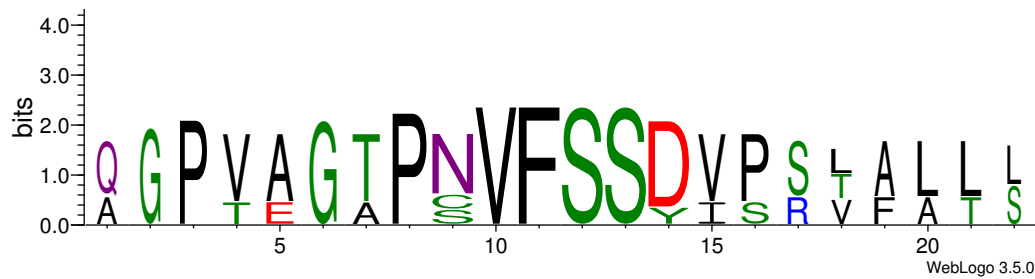
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2544	LSSNLVILPGQDLQ	____LSSNLVILPGQDLQ_	3	1	2
2544	VKLSSNLVILPGQD	___VKLSSNLVILPGQD___	1	0	1
2544	SSNLVILPGQDLQY	____SSNLVILPGQDLQY	1	0	1
2544	KLSSNLVILPGQDL	____KLSSNLVILPGQDL__	1	0	1
2544	GGDVKLSSNLVILP	GGDVKLSSNLVILP_____	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
GGDVKLSSNLVILP_____	X89184	MV-H-CAM-70	505	14aaG4S	mRNA_30cpc_Organoid_MD114	1
GGDVKLSSNLVILP_____	X89184	MV-H-Edmonston	505	14aaG4S	mRNA_30cpc_Organoid_MD114	1
GGDVKLSSNLVILP_____	X89184	MV-H-Halle	505	14aaG4S	mRNA_30cpc_Organoid_MD114	1
GGDVKLSSNLVILP_____	X89184	MV-H-MVi-Illinois.USA-50.99	505	14aaG4S	mRNA_30cpc_Organoid_MD114	1
GGDVKLSSNLVILP_____	X89184	MV-H-MVi-Texas.USA-28.99	505	14aaG4S	mRNA_30cpc_Organoid_MD114	1
____KLSSNLVILPGQDL__	X4231	MV-H-CAM-70	509	14aa	mRNA_3000cpc_Organoid_MD101	1
____KLSSNLVILPGQDL__	X4231	MV-H-Edmonston	509	14aa	mRNA_3000cpc_Organoid_MD101	1
____KLSSNLVILPGQDL__	X4231	MV-H-Halle	509	14aa	mRNA_3000cpc_Organoid_MD101	1
____KLSSNLVILPGQDL__	X4231	MV-H-MVi-Illinois.USA-50.99	509	14aa	mRNA_3000cpc_Organoid_MD101	1
____KLSSNLVILPGQDL__	X4231	MV-H-MVi-Texas.USA-28.99	509	14aa	mRNA_3000cpc_Organoid_MD101	1
____LSSNLVILPGQDLQ_	X29170	MV-H-CAM-70	510	14aa	mRNA_30cpc_Organoid_MD114	1
____LSSNLVILPGQDLQ_	X29170	MV-H-Edmonston	510	14aa	mRNA_30cpc_Organoid_MD114	1
____LSSNLVILPGQDLQ_	X29170	MV-H-Halle	510	14aa	mRNA_30cpc_Organoid_MD114	1
____LSSNLVILPGQDLQ_	X29170	MV-H-MVi-Illinois.USA-50.99	510	14aa	mRNA_30cpc_Organoid_MD114	1
____LSSNLVILPGQDLQ_	X29170	MV-H-MVi-Texas.USA-28.99	510	14aa	mRNA_30cpc_Organoid_MD114	1
____LSSNLVILPGQDLQ_	X29170	MV-H-CAM-70	510	14aa	mRNA_3000cpc_Organoid_MD101	2
____LSSNLVILPGQDLQ_	X29170	MV-H-Edmonston	510	14aa	mRNA_3000cpc_Organoid_MD101	2
____LSSNLVILPGQDLQ_	X29170	MV-H-Halle	510	14aa	mRNA_3000cpc_Organoid_MD101	2
____LSSNLVILPGQDLQ_	X29170	MV-H-MVi-Illinois.USA-50.99	510	14aa	mRNA_3000cpc_Organoid_MD101	2
____LSSNLVILPGQDLQ_	X29170	MV-H-MVi-Texas.USA-28.99	510	14aa	mRNA_3000cpc_Organoid_MD101	2
____SSNLVILPGQDLQY	X10561	MV-H-CAM-70	511	14aa	mRNA_3000cpc_Organoid_MD101	1
____SSNLVILPGQDLQY	X10561	MV-H-Edmonston	511	14aa	mRNA_3000cpc_Organoid_MD101	1
____SSNLVILPGQDLQY	X10561	MV-H-Halle	511	14aa	mRNA_3000cpc_Organoid_MD101	1
____SSNLVILPGQDLQY	X10561	MV-H-MVi-Illinois.USA-50.99	511	14aa	mRNA_3000cpc_Organoid_MD101	1
____SSNLVILPGQDLQY	X10561	MV-H-MVi-Texas.USA-28.99	511	14aa	mRNA_3000cpc_Organoid_MD101	1
___VKLSSNLVILPGQD___	X65861	MV-H-CAM-70	508	14aa	mRNA_3000cpc_Organoid_MD101	1
___VKLSSNLVILPGQD___	X65861	MV-H-Edmonston	508	14aa	mRNA_3000cpc_Organoid_MD101	1
___VKLSSNLVILPGQD___	X65861	MV-H-Halle	508	14aa	mRNA_3000cpc_Organoid_MD101	1
___VKLSSNLVILPGQD___	X65861	MV-H-MVi-Illinois.USA-50.99	508	14aa	mRNA_3000cpc_Organoid_MD101	1
___VKLSSNLVILPGQD___	X65861	MV-H-MVi-Texas.USA-28.99	508	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2826

Peptide QGPVAGTPNVFSSD from HSV-2-pUL22 with cluster number 2826

Peptide QGPVAGTPNVFSSD from HSV-2-pUL22 with cluster number 2826



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2826	QGPVAGTPNVFSSD	6	4	2

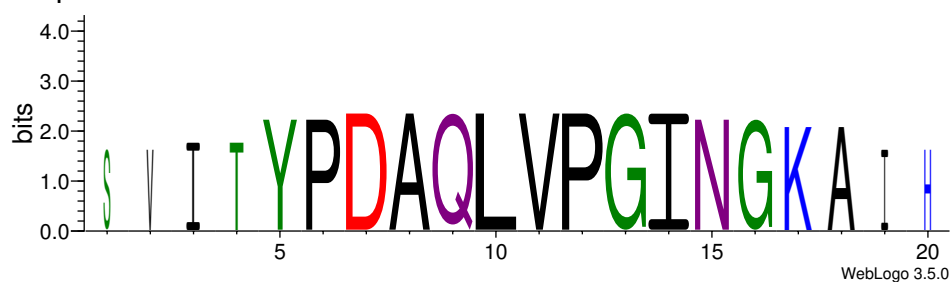
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2826	QGPVAGTPNVFSSD	QGPVAGTPNVFSSD_____	2	1	1
2826	PVAGTPNVFSSDVP	__PVAGTPNVFSSDVP_____	1	0	1
2826	PNVFSSDVPSTALL	_____PNVFSSDVPSTALL_	1	1	0
2826	CVFSSYISRLFATS	_____CVFSSYISRLFATS	1	1	0
2826	AGPTEGAPSVFSSDVPSTALL	AGPTEGAPSVFSSDVPSTALL	1	1	0

alignment	LUTnrs	GeneName	start	structure	Group	Score
AGPTEGAPSVFSSDVPSTALL	X44043	HSV-1-pUL22	760	22aa	mRNA_30cpc_Organoid_MD114	1
_____CVFSSYISRLFATS	X73281	HSV-1-pUL16	343	14aa	mRNA_30cpc_Organoid_MD114	1
_____PNVFSSDVPSTALL_	X20600	HSV-1-pUL22	767	14aa	mRNA_30cpc_Organoid_MD114	1
_____PNVFSSDVPSTALL_	X20600	HSV-2-pUL22	767	14aa	mRNA_30cpc_Organoid_MD114	1
__PVAGTPNVFSSDVP_____	X24297	HSV-2-pUL22	762	14aa	mRNA_3000cpc_Organoid_MD101	1
QGPVAGTPNVFSSD_____	X50456	HSV-2-pUL22	760	14aaA5	mRNA_30cpc_Organoid_MD114	1
QGPVAGTPNVFSSD_____	X50456	HSV-2-pUL22	760	14aaA5	mRNA_3000cpc_Organoid_MD101	1

[1] 2806

Peptide YPDAQLVPGINGKA from TeNT-Hc with cluster number 2806

‡ Peptide YPDAQLVPGINGKA from TeNT-Hc with cluster number 2806



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2806	YPDAQLVPGINGKA	6	3	3

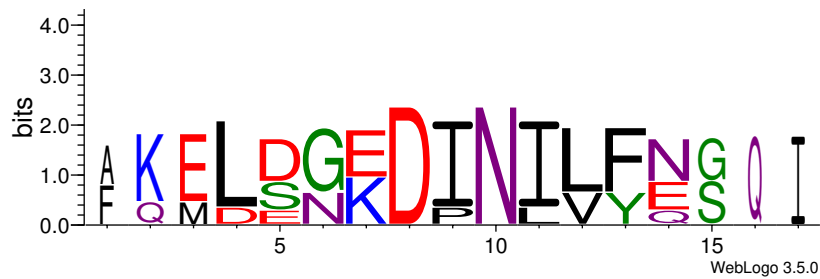
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2806	YPDAQLVPGINGKA	____YPDAQLVPGINGKA__	3	2	1
2806	SVITYPDAQLVPGI	SVITYPDAQLVPGI_____	1	1	0
2806	ITYPDAQLVPGING	__ITYPDAQLVPGING_____	1	0	1
2806	DAQLVPGINGKAIH	______DAQLVPGINGKAIH	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
______DAQLVPGINGKAIH	X35832	TeNT-Hc	48	14aa	mRNA_3000cpc_Organoid_MD101	1
__ITYPDAQLVPGING_____	X14700	TeNT-Hc	44	14aa	mRNA_3000cpc_Organoid_MD101	1
SVITYPDAQLVPGI_____	X13484	TeNT-Hc	42	14aa	mRNA_30cpc_Organoid_MD114	1
____YPDAQLVPGINGKA__	X70921	TeNT-Hc	46	14aa	mRNA_30cpc_Organoid_MD114	1
____YPDAQLVPGINGKA__	X59590	TeNT-Hc	46	14aaA5	mRNA_30cpc_Organoid_MD114	1
____YPDAQLVPGINGKA__	X59590	TeNT-Hc	46	14aaA5	mRNA_3000cpc_Organoid_MD101	1

[1] 2805

Peptide KELDGKDINILFNS from BoNT-C-Hc with cluster number 2805

Peptide KELDGKDINILFNS from BoNT-C-Hc with cluster number 2805



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2805	KELDGKDINILFNS	6	4	2

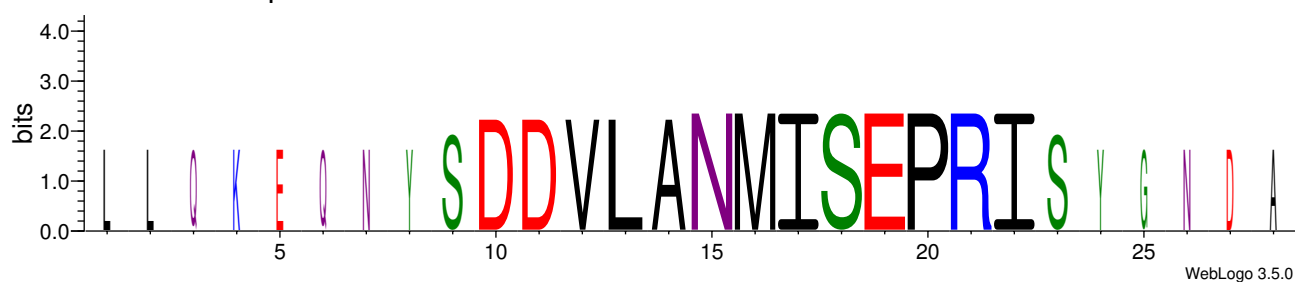
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2805	LSNEDINIVYEGQI	___LSNEDINIVYEGQI	2	0	2
2805	KELDGKDINILFNS	_KELDGKDINILFNS__	2	2	0
2805	FQMDEGEDPNLLFQ	FQMDEGEDPNLLFQ___	1	1	0
2805	AKELDGKDINILFN	AKELDGKDINILFN___	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
AKELDGKDINILFN___	X40071	BoNT-C-Hc	217	14aa	mRNA_30cpc_Organoid_MD114	1
FQMDEGEDPNLLFQ___	X75093	RV-P	81	14aa	mRNA_30cpc_Organoid_MD114	1
_KELDGKDINILFNS__	X4632	BoNT-C-Hc	218	14aa	mRNA_30cpc_Organoid_MD114	2
___LSNEDINIVYEGQI	X29122	BoNT-D-Hc	204	14aa	mRNA_3000cpc_Organoid_MD101	2

[1] 2800

Peptide DDVLANMISEPRIS from APP with cluster number 2800

Peptide DDVLANMISEPRIS from APP with cluster number 2800



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2800	DDVLANMISEPRIS	6	2	4

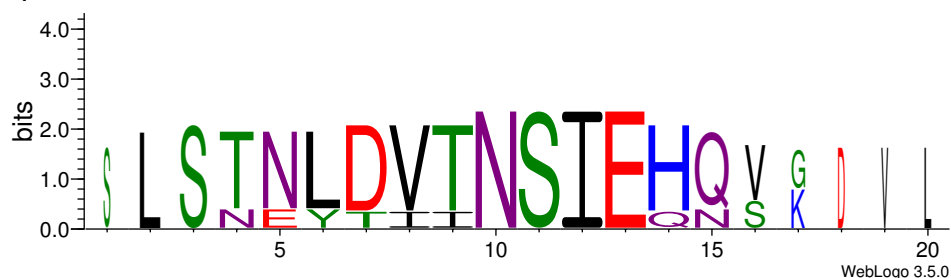
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2800	SDDVLANMISEPRI	-----SDDVLANMISEPRI-----	2	1	1
2800	DDVLANMISEPRIS	-----DDVLANMISEPRIS-----	2	1	1
2800	NMISEPRISYGND	-----NMISEPRISYGND-----	1	0	1
2800	LLQKEQNYSDVLNMISEPRI	LLQKEQNYSDVLNMISEPRI-----	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
-----DDVLNMISEPRIS-----	X54506	APP	574	14aaA5	mRNA_30cpc_Organoid_MD114	1
-----DDVLNMISEPRIS-----	X35610	APP	574	14aa	mRNA_3000cpc_Organoid_MD101	1
LLQKEQNYSDVLNMISEPRI-----	X30879	APP	565	22aa	mRNA_3000cpc_Organoid_MD101	1
-----NMISEPRISYGND	X963	APP	579	14aa	mRNA_3000cpc_Organoid_MD101	1
-----SDDVLNMISEPRI-----	X12510	APP	573	14aa	mRNA_30cpc_Organoid_MD114	1
-----SDDVLNMISEPRI-----	X12510	APP	573	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2799

Peptide LSTNLDVTNSIEHQ from MV-H-CAM-70 with cluster number 2799

Peptide LSTNLDVTNSIEHQ from MV-H-CAM-70 with cluster number 2799



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2799	LSTNLDVTNSIEHQ	6	2	4

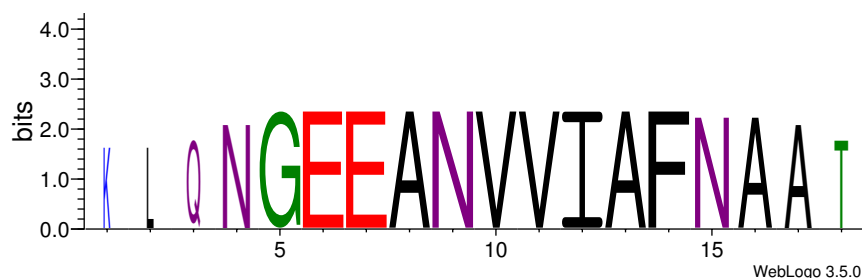
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2799	LSTNLDVTNSIEHQ	_LSTNLDVTNSIEHQ_	2	1	1
2799	STNLDVTNSIEHQV	_STNLDVTNSIEHQV_	1	0	1
2799	SLSTNLDVTNSIEH	SLSTNLDVTNSIEH_	1	0	1
2799	NEYTIINSIEQNSG	__NEYTIINSIEQNSG__	1	1	0
2799	DVTNSIEHQVKDVL	____DVTNSIEHQVKDVL	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
____DVTNSIEHQVKDVL	X36230	MV-H-CAM-70	79	14aa	mRNA_3000cpc_Organoid_MD101	1
____DVTNSIEHQVKDVL	X36230	MV-H-Edmonston	79	14aa	mRNA_3000cpc_Organoid_MD101	1
____DVTNSIEHQVKDVL	X36230	MV-H-Halle	79	14aa	mRNA_3000cpc_Organoid_MD101	1
____DVTNSIEHQVKDVL	X36230	MV-H-MVi-Illinois.USA-50.99	79	14aa	mRNA_3000cpc_Organoid_MD101	1
____DVTNSIEHQVKDVL	X36230	MV-H-MVi-Texas.USA-28.99	79	14aa	mRNA_3000cpc_Organoid_MD101	1
LSTNLDVTNSIEHQ	X29145	MV-H-CAM-70	74	14aa	mRNA_30cpc_Organoid_MD114	1
LSTNLDVTNSIEHQ	X29145	MV-H-Edmonston	74	14aa	mRNA_30cpc_Organoid_MD114	1
LSTNLDVTNSIEHQ	X29145	MV-H-Halle	74	14aa	mRNA_30cpc_Organoid_MD114	1
LSTNLDVTNSIEHQ	X29145	MV-H-MVi-Illinois.USA-50.99	74	14aa	mRNA_30cpc_Organoid_MD114	1
LSTNLDVTNSIEHQ	X29145	MV-H-MVi-Texas.USA-28.99	74	14aa	mRNA_30cpc_Organoid_MD114	1
LSTNLDVTNSIEHQ	X29145	MV-H-CAM-70	74	14aa	mRNA_3000cpc_Organoid_MD101	1
LSTNLDVTNSIEHQ	X29145	MV-H-Edmonston	74	14aa	mRNA_3000cpc_Organoid_MD101	1
LSTNLDVTNSIEHQ	X29145	MV-H-Halle	74	14aa	mRNA_3000cpc_Organoid_MD101	1
LSTNLDVTNSIEHQ	X29145	MV-H-MVi-Illinois.USA-50.99	74	14aa	mRNA_3000cpc_Organoid_MD101	1
LSTNLDVTNSIEHQ	X29145	MV-H-MVi-Texas.USA-28.99	74	14aa	mRNA_3000cpc_Organoid_MD101	1
__NEYTIINSIEQNSG__	X77063	BoNT-A-Hc	127	14aaG4S	mRNA_30cpc_Organoid_MD114	1
__NEYTIINSIEQNSG__	X77063	BoNT-D-Hc	85	14aaG4S	mRNA_30cpc_Organoid_MD114	1
SLSTNLDVTNSIEH	X48400	MV-H-CAM-70	73	14aaA5	mRNA_3000cpc_Organoid_MD101	1
SLSTNLDVTNSIEH	X48400	MV-H-Edmonston	73	14aaA5	mRNA_3000cpc_Organoid_MD101	1
SLSTNLDVTNSIEH	X48400	MV-H-Halle	73	14aaA5	mRNA_3000cpc_Organoid_MD101	1
SLSTNLDVTNSIEH	X48400	MV-H-MVi-Illinois.USA-50.99	73	14aaA5	mRNA_3000cpc_Organoid_MD101	1
SLSTNLDVTNSIEH	X48400	MV-H-MVi-Texas.USA-28.99	73	14aaA5	mRNA_3000cpc_Organoid_MD101	1
STNLDVTNSIEHQV	X10229	MV-H-CAM-70	75	14aa	mRNA_3000cpc_Organoid_MD101	1
STNLDVTNSIEHQV	X10229	MV-H-Edmonston	75	14aa	mRNA_3000cpc_Organoid_MD101	1
STNLDVTNSIEHQV	X10229	MV-H-Halle	75	14aa	mRNA_3000cpc_Organoid_MD101	1
STNLDVTNSIEHQV	X10229	MV-H-MVi-Illinois.USA-50.99	75	14aa	mRNA_3000cpc_Organoid_MD101	1
STNLDVTNSIEHQV	X10229	MV-H-MVi-Texas.USA-28.99	75	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2797

Peptide GEEANVVIAFNAAT from PSA-Chain-C with cluster number 2797

Peptide GEEANVVIAFNAAT from PSA-Chain-C with cluster number



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2797	GEEANVVIAFNAAT	6	1	5

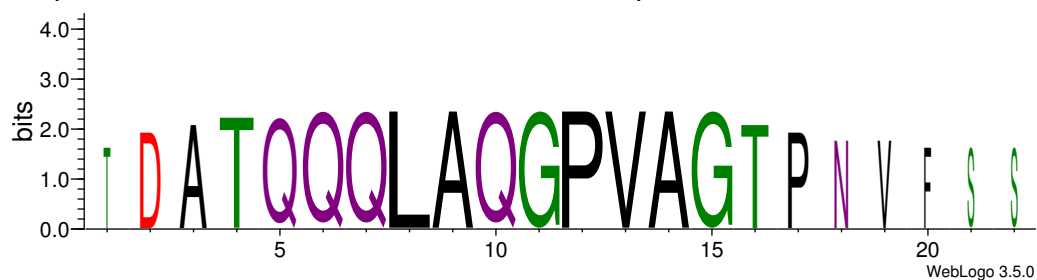
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2797	NGEEANVVIAFNAA	___NGEEANVVIAFNAA_	2	0	2
2797	GEEANVVIAFNAAT	____GEEANVVIAFNAAT	2	1	1
2797	QNGEEANVVIAFNA	__QNGEEANVVIAFNA__	1	0	1
2797	KLQNGEEANVVIAF	KLQNGEEANVVIAF_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
____GEEANVVIAFNAAT	X63883	PSA-Chain-C	157	14aa	mRNA_30cpc_Organoid_MD114	1
____GEEANVVIAFNAAT	X57709	PSA-Chain-C	157	14aaA5	mRNA_3000cpc_Organoid_MD101	1
KLQNGEEANVVIAF_____	X4253	PSA-Chain-C	153	14aa	mRNA_3000cpc_Organoid_MD101	1
___NGEEANVVIAFNAA_	X2412	PSA-Chain-C	156	14aa	mRNA_3000cpc_Organoid_MD101	2
__QNGEEANVVIAFNA__	X18495	PSA-Chain-C	155	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2775

Peptide DATQQQLAQGPVAG from HSV-2-pUL22 with cluster number 2775

† Peptide DATQQQLAQGPVAG from HSV-2-pUL22 with cluster number 2775



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2775	DATQQQLAQGPVAG	6	1	5

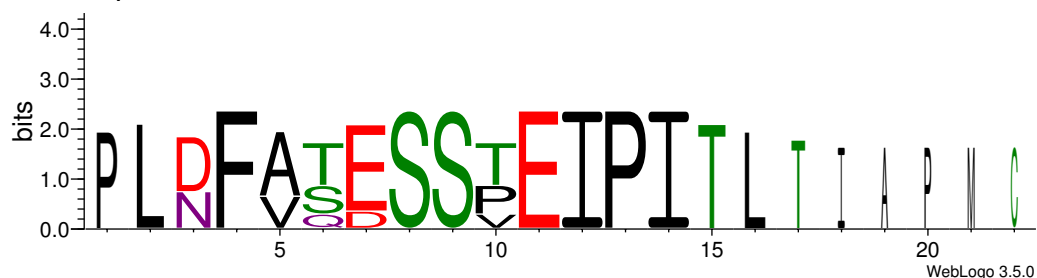
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2775	DATQQQLAQGPVAG	_DATQQQLAQGPVAG_	2	0	2
2775	TQQQLAQGPVAGTP	__TQQQLAQGPVAGTP	1	0	1
2775	TDATQQQLAQGPVAGTPNVFSS	TDATQQQLAQGPVAGTPNVFSS	1	0	1
2775	QQLAQGPVAGTPNV	____QQLAQGPVAGTPNV__	1	0	1
2775	ATQQQLAQGPVAGT	__ATQQQLAQGPVAGT_	1	1	0

alignment	LUTnrs	GeneName	start	structure	Group	Score
__ATQQQLAQGPVAGT_	X40235	HSV-2-pUL22	753	14aa	mRNA_30cpc_Organoid_MD114	1
DATQQQLAQGPVAG	X35790	HSV-2-pUL22	752	14aa	mRNA_3000cpc_Organoid_MD101	2
____QQLAQGPVAGTPNV__	X19148	HSV-2-pUL22	756	14aa	mRNA_3000cpc_Organoid_MD101	1
TDATQQQLAQGPVAGTPNVFSS	X8153	HSV-2-pUL22	751	22aa	mRNA_3000cpc_Organoid_MD101	1
__TQQQLAQGPVAGTP	X7009	HSV-2-pUL22	754	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2753

Peptide FATESSTEIPITLT from PV1-VP2 with cluster number 2753

Peptide FATESSTEIPITLT from PV1-VP2 with cluster number 2753



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2753	FATESSTEIPITLT	6	3	3

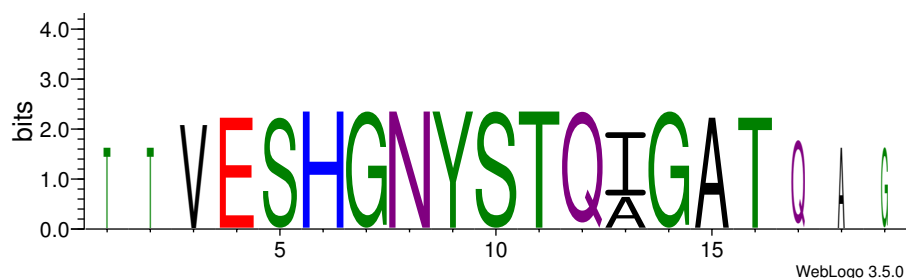
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2753	PLDFATESSTEIPITLTIAPMC	PLDFATESSTEIPITLTIAPMC	1	1	0
2753	PLDFATESSTEIPI	PLDFATESSTEIPI_	1	1	0
2753	PLDFAQDSSVEIPI	PLDFAQDSSVEIPI_	1	1	0
2753	NFVSESSPEIPITL	_NFVSESSPEIPITL_	1	0	1
2753	LNFVSESSPEIPIT	_LNFVSESSPEIPIT_	1	0	1
2753	FATESSTEIPITLT	__FATESSTEIPITLT__	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
__FATESSTEIPITLT__	X75762	PV1-VP2	238	14aa	mRNA_3000cpc_Organoid_MD101	1
__FATESSTEIPITLT__	X75762	PV2-VP2	238	14aa	mRNA_3000cpc_Organoid_MD101	1
LNFVSESSPEIPIT	X28523	PV1-VP2	236	14aa	mRNA_3000cpc_Organoid_MD101	1
NFVSESSPEIPITL	X2982	PV1-VP2	237	14aa	mRNA_3000cpc_Organoid_MD101	1
PLDFAQDSSVEIPI_	X82418	PV3-VP1	304	14aaG4S	mRNA_30cpc_Organoid_MD114	1
PLDFATESSTEIPI_	X22728	PV2-VP2	235	14aa	mRNA_30cpc_Organoid_MD114	1
PLDFATESSTEIPITLTIAPMC	X22727	PV1-VP2	235	22aa	mRNA_30cpc_Organoid_MD114	1
PLDFATESSTEIPITLTIAPMC	X22727	PV2-VP2	235	22aa	mRNA_30cpc_Organoid_MD114	1

[1] 2715

Peptide VESHGNYSTQIGAT from WNV-E-AZ10-918 with cluster number 2715

ide VESHGNYSTQIGAT from WNV-E-AZ10-918 with cluster numc



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2715	VESHGNYSTQIGAT	6	2	4

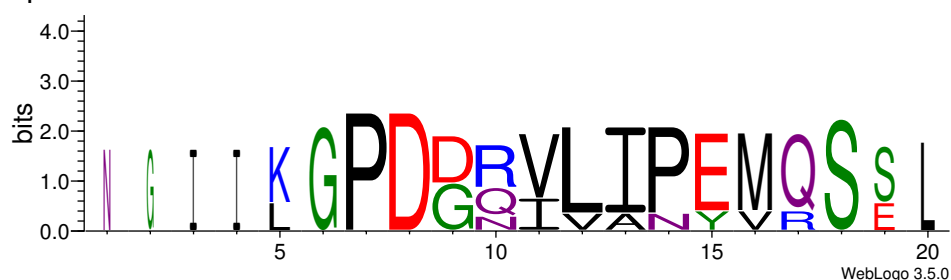
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2715	VESHGNYSTQIGAT	__VESHGNYSTQIGAT__	2	0	2
2715	VESHGNYSTQAGAT	__VESHGNYSTQAGAT__	1	0	1
2715	TTVESHGNYSTQAG	TTVESHGNYSTQAG_____	1	1	0
2715	HGNYSTQIGATQAG	_____HGNYSTQIGATQAG	1	1	0
2715	ESHGNYSTQIGATQ	___ESHGNYSTQIGATQ__	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
___ESHGNYSTQIGATQ__	X37279	WNV-E-AZ10-918	150	14aa	mRNA_3000cpc_Organoid_MD101	1
___ESHGNYSTQIGATQ__	X37279	WNV-E-H-442	150	14aa	mRNA_3000cpc_Organoid_MD101	1
_____HGNYSTQIGATQAG	X18087	WNV-E-AZ10-918	152	14aa	mRNA_30cpc_Organoid_MD114	1
_____HGNYSTQIGATQAG	X18087	WNV-E-H-442	152	14aa	mRNA_30cpc_Organoid_MD114	1
TTVESHGNYSTQAG_____	X6160	WNV-E-AZ10-918	147	14aa	mRNA_30cpc_Organoid_MD114	1
TTVESHGNYSTQAG_____	X6160	WNV-E-H-442	147	14aa	mRNA_30cpc_Organoid_MD114	1
__VESHGNYSTQAGAT__	X68262	WNV-E-AZ10-918	149	14aa	mRNA_3000cpc_Organoid_MD101	1
__VESHGNYSTQAGAT__	X68262	WNV-E-H-442	149	14aa	mRNA_3000cpc_Organoid_MD101	1
__VESHGNYSTQIGAT__	X68261	WNV-E-AZ10-918	149	14aa	mRNA_3000cpc_Organoid_MD101	2
__VESHGNYSTQIGAT__	X68261	WNV-E-H-442	149	14aa	mRNA_3000cpc_Organoid_MD101	2

[1] 2671

Peptide KGPDPGQILIPEMQS from RV-G-Mocala with cluster number 2671

Peptide KGPDPGQILIPEMQS from RV-G-Mocala with cluster number 2



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2671	KGPDPGQILIPEMQS	6	4	2

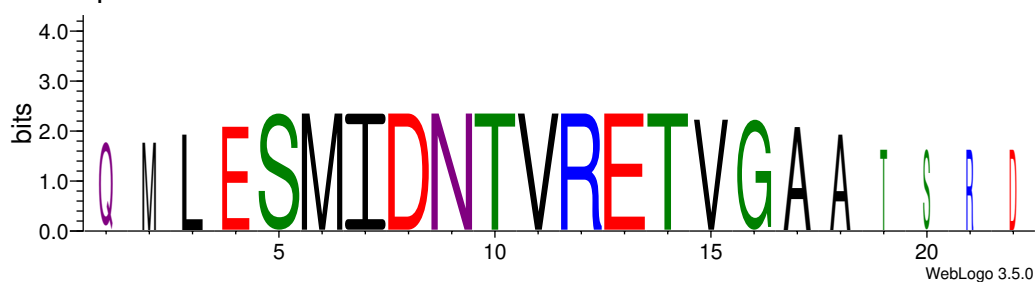
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2671	KGPDPGQILIPEMQS	___KGPDPGQILIPEMQS__	2	2	0
2671	PDDRVVANYVRSEL	_____PDDRVVANYVRSEL	1	0	1
2671	PDDRVLIPEMQSSL	_____PDDRVLIPEMQSSL	1	1	0
2671	NGIILGPDDRVLIP	NGIILGPDDRVLIP_____	1	1	0
2671	GPDGNVLIPEMQSS	_____GPDGNVLIPEMQSS_	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
_____GPDGNVLIPEMQSS_	X62680	RV-G-B19	366	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GPDGNVLIPEMQSS_	X62680	RV-G-B2c	366	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GPDGNVLIPEMQSS_	X62680	RV-G-N2c	385	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GPDGNVLIPEMQSS_	X62680	RV-G-PV	366	14aa	mRNA_3000cpc_Organoid_MD101	1
___KGPDPGQILIPEMQS__	X4926	RV-G-Mocala	365	14aa	mRNA_30cpc_Organoid_MD114	2
NGIILGPDDRVLIP_____	X45853	RV-G-B19	361	14aaA5	mRNA_30cpc_Organoid_MD114	1
NGIILGPDDRVLIP_____	X45853	RV-G-B2c	361	14aaA5	mRNA_30cpc_Organoid_MD114	1
NGIILGPDDRVLIP_____	X45853	RV-G-N2c	380	14aaA5	mRNA_30cpc_Organoid_MD114	1
NGIILGPDDRVLIP_____	X45853	RV-G-PV	361	14aaA5	mRNA_30cpc_Organoid_MD114	1
_____PDDRVLIPEMQSSL	X82482	RV-G-B19	367	14aaG4S	mRNA_30cpc_Organoid_MD114	1
_____PDDRVLIPEMQSSL	X82482	RV-G-B2c	367	14aaG4S	mRNA_30cpc_Organoid_MD114	1
_____PDDRVLIPEMQSSL	X82482	RV-G-N2c	386	14aaG4S	mRNA_30cpc_Organoid_MD114	1
_____PDDRVLIPEMQSSL	X82482	RV-G-PV	367	14aaG4S	mRNA_30cpc_Organoid_MD114	1
_____PDDRVVANYVRSEL	X22978	HSV-1-pUL21	59	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2620

Peptide SMIDNTVRETVGAA from PV1-VP1 with cluster number 2620

Peptide SMIDNTVRETVGAA from PV1-VP1 with cluster number 2620



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2620	SMIDNTVRETVGAA	6	1	5

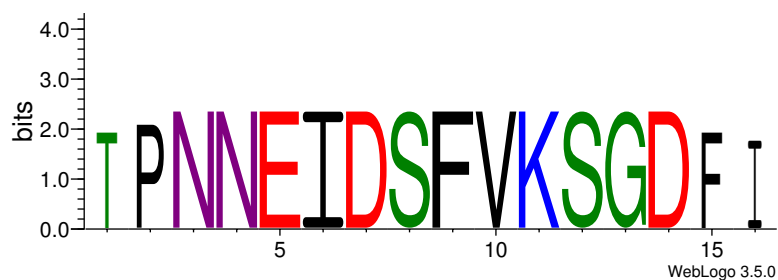
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2620	SMIDNTVRETVGAA	___SMIDNTVRETVGAA___	2	1	1
2620	QMLES MIDNTVRETVGAATSRD	QMLES MIDNTVRETVGAATSRD	1	0	1
2620	QMLES MIDNTVRET	QMLES MIDNTVRET_____	1	0	1
2620	LES MIDNTVRETVG	__LES MIDNTVRETVG_____	1	0	1
2620	ES MIDNTVRETVGA	___ES MIDNTVRETVGA_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
___ES MIDNTVRETVGA_____	X37273	PV1-VP1	7	14aa	mRNA_3000cpc_Organoid_MD101	1
__LES MIDNTVRETVG_____	X31535	PV1-VP1	6	14aa	mRNA_3000cpc_Organoid_MD101	1
QMLES MIDNTVRET_____	X50196	PV1-VP1	4	14aaA5	mRNA_3000cpc_Organoid_MD101	1
QMLES MIDNTVRETVGAATSRD	X19027	PV1-VP1	4	22aa	mRNA_3000cpc_Organoid_MD101	1
___SMIDNTVRETVGAA_____	X11154	PV1-VP1	8	14aa	mRNA_30cpc_Organoid_MD114	1
___SMIDNTVRETVGAA_____	X11154	PV1-VP1	8	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2603

Peptide TPNNEIDSFVKSGD from TeNT-Hc with cluster number 2603

ide TPNNEIDSFVKSGD from TeNT-Hc with cluster numb



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2603	TPNNEIDSFVKSGD	6	3	3

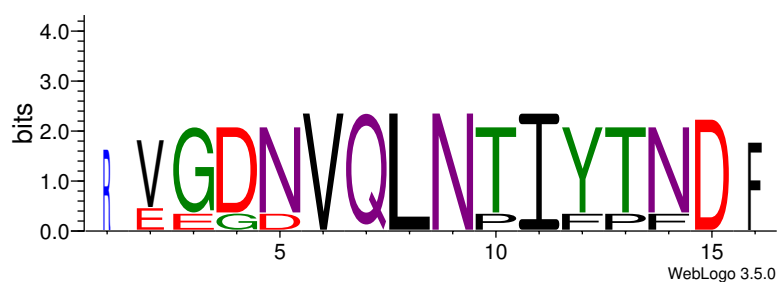
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2603	TPNNEIDSFVKSGD	TPNNEIDSFVKSGD__	3	2	1
2603	NNEIDSFVKSGDFI	__NNEIDSFVKSGDFI	2	1	1
2603	PNNEIDSFVKSGDF	_PNNEIDSFVKSGDF_	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
__NNEIDSFVKSGDFI	X169	TeNT-Hc	320	14aa	mRNA_30cpc_Organoid_MD114	1
__NNEIDSFVKSGDFI	X169	TeNT-Hc	320	14aa	mRNA_3000cpc_Organoid_MD101	1
PNNEIDSFVKSGDF	X20476	TeNT-Hc	319	14aa	mRNA_3000cpc_Organoid_MD101	1
TPNNEIDSFVKSGD__	X7129	TeNT-Hc	318	14aa	mRNA_30cpc_Organoid_MD114	2
TPNNEIDSFVKSGD__	X7129	TeNT-Hc	318	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2584

Peptide GDNVQLNTIYTND F from BoNT-D-Hc with cluster number 2584

le GDNVQLNTIYTND F from BoNT-D-Hc with cluster num



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2584	GDNVQLNTIYTND F	6	4	2

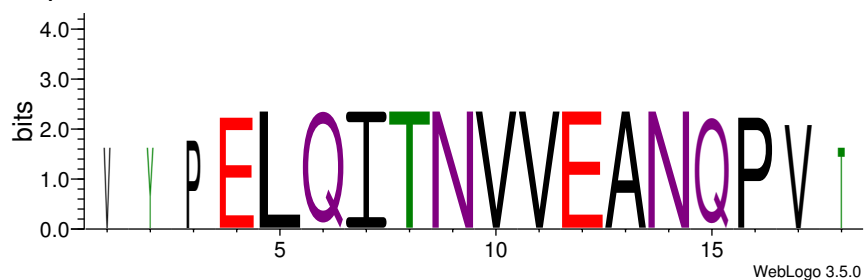
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2584	VGDNVQLNTIYTND	_VGDNVQLNTIYTND_	2	1	1
2584	GDNVQLNTIYTND F	__GDNVQLNTIYTND F	2	1	1
2584	RVGDNVQLNTIYTN	RVGDNVQLNTIYTN__	1	1	0
2584	EEGDVQLNPIFPFD	_EEGDVQLNPIFPFD_	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
EEGDVQLNPIFPFD	X38657	BoNT-C-Hc	35	14aa	mRNA_30cpc_Organoid_MD114	1
__GDNVQLNTIYTND F	X63483	BoNT-D-Hc	29	14aa	mRNA_30cpc_Organoid_MD114	1
__GDNVQLNTIYTND F	X63483	BoNT-D-Hc	29	14aa	mRNA_3000cpc_Organoid_MD101	1
RVGDNVQLNTIYTN__	X27836	BoNT-D-Hc	27	14aa	mRNA_30cpc_Organoid_MD114	1
VGDNVQLNTIYTND	X69090	BoNT-D-Hc	28	14aa	mRNA_30cpc_Organoid_MD114	1
VGDNVQLNTIYTND	X69090	BoNT-D-Hc	28	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2550

Peptide ELQITNVVEANQPV from APP with cluster number 2550

Peptide ELQITNVVEANQPV from APP with cluster number 25



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2550	ELQITNVVEANQPV	6	5	1

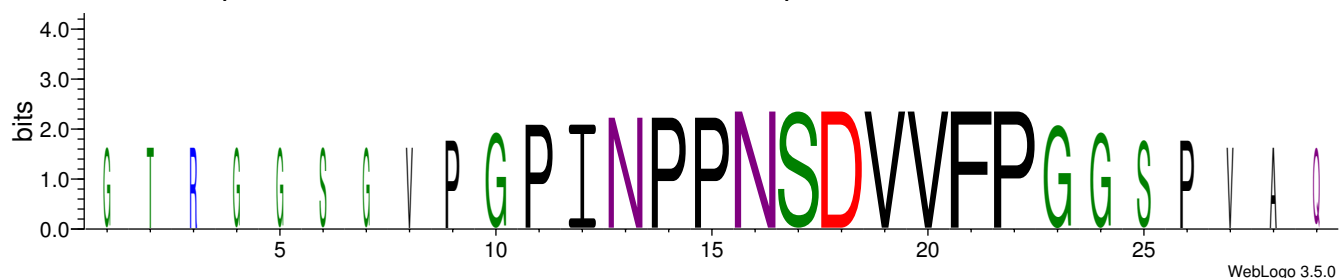
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2550	ELQITNVVEANQPV	___ELQITNVVEANQPV_	3	2	1
2550	VYPELQITNVVEAN	VYPELQITNVVEAN_____	1	1	0
2550	PELQITNVVEANQP	__PELQITNVVEANQP__	1	1	0
2550	LQITNVVEANQPVT	____LQITNVVEANQPVT	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
___ELQITNVVEANQPV_	X38162	APP	79	14aa	mRNA_30cpc_Organoid_MD114	1
___ELQITNVVEANQPV_	X86438	APP	79	14aaG4S	mRNA_30cpc_Organoid_MD114	1
___ELQITNVVEANQPV_	X38162	APP	79	14aa	mRNA_3000cpc_Organoid_MD101	1
____LQITNVVEANQPVT	X29973	APP	80	14aa	mRNA_30cpc_Organoid_MD114	1
__PELQITNVVEANQP__	X23144	APP	78	14aa	mRNA_30cpc_Organoid_MD114	1
VYPELQITNVVEAN_____	X59232	APP	76	14aaA5	mRNA_30cpc_Organoid_MD114	1

[1] 2412

Peptide GPINPPNSDVVFPG from HSV-2-pUS4 with cluster number 2412

Peptide GPINPPNSDVVFPG from HSV-2-pUS4 with cluster number 2412



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2412	GPINPPNSDVVFPG	6	4	2

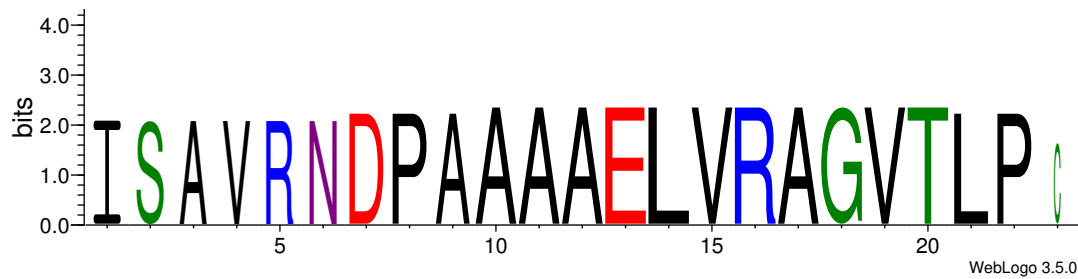
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2412	PINPPNSDVVFPGG	-----PINPPNSDVVFPGG-----	1	1	0
2412	PGPINPPNSDVVF	-----PGPINPPNSDVVF-----	1	1	0
2412	NSDVVFPGGSPVAQ	-----NSDVVFPGGSPVAQ-----	1	1	0
2412	NPPNSDVVFPGGSP	-----NPPNSDVVFPGGSP-----	1	0	1
2412	GTRGGSGVPGPINPPNSDVVF	GTRGGSGVPGPINPPNSDVVF-----	1	0	1
2412	GPINPPNSDVVFPG	-----GPINPPNSDVVFPG-----	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
-----GPINPPNSDVVFPG-----	X57364	HSV-2-pUS4	28	14aaA5	mRNA_30cpc_Organoid_MD114	1
GTRGGSGVPGPINPPNSDVVF-----	X61571	HSV-2-pUS4	19	22aa	mRNA_3000cpc_Organoid_MD101	1
-----NPPNSDVVFPGGSP-----	X45580	HSV-2-pUS4	31	14aaA5	mRNA_3000cpc_Organoid_MD101	1
-----NSDVVFPGGSPVAQ-----	X45446	HSV-2-pUS4	34	14aaA5	mRNA_30cpc_Organoid_MD114	1
-----PGPINPPNSDVVF-----	X23843	HSV-2-pUS4	27	14aa	mRNA_30cpc_Organoid_MD114	1
-----PINPPNSDVVFPG-----	X21380	HSV-2-pUS4	29	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2263

Peptide ISAVRNDPAAAAELVRAGVTLP from HSV-1-pUL36 with cluster number 2263

Peptide ISAVRNDPAAAAELVRAGVTLP from HSV-1-pUL36 with cluster number



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2263	ISAVRNDPAAAAELVRAGVTLP	6	2	4

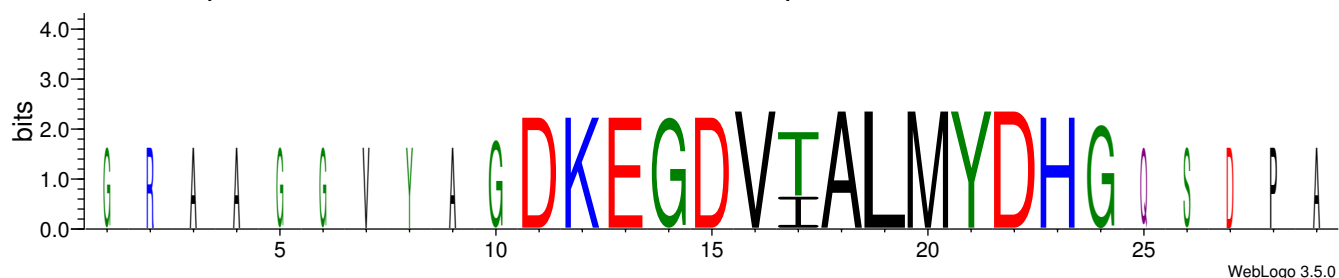
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2263	ISAVRNDPAAAAELVRAGVTLP	ISAVRNDPAAAAELVRAGVTLP_	4	1	3
2263	DPAAAAELVRAGVT	-----DPAAAAELVRAGVT---	1	0	1
2263	AAAELVRAGVTLPC	-----AAAELVRAGVTLPC	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
-----AAAELVRAGVTLPC	X43771	HSV-1-pUL36	1534	14aa	mRNA_30cpc_Organoid_MD114	1
-----DPAAAAELVRAGVT---	X34988	HSV-1-pUL36	1531	14aa	mRNA_3000cpc_Organoid_MD101	1
ISAVRNDPAAAAELVRAGVTLP_	X14875	HSV-1-pUL36	1525	22aa	mRNA_30cpc_Organoid_MD114	1
ISAVRNDPAAAAELVRAGVTLP_	X14875	HSV-1-pUL36	1525	22aa	mRNA_3000cpc_Organoid_MD101	3

[1] 2156

Peptide DKEGDVTALMYDHG from HSV-1-pUL19 with cluster number 2156

Peptide DKEGDVTALMYDHG from HSV-1-pUL19 with cluster number 2156



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2156	DKEGDVTALMYDHG	6	3	3

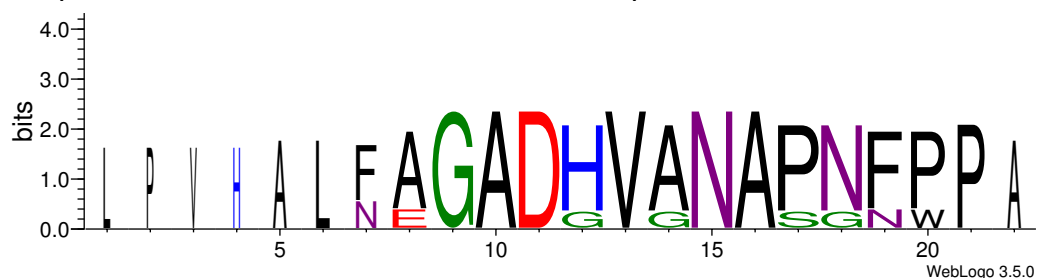
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2156	DKEGDVTALMYDHG	-----DKEGDVTALMYDHG-----	3	1	2
2156	VIALMYDHGQSDPA	-----VIALMYDHGQSDPA-----	1	1	0
2156	GRAAGGVYAGDKEGDVIALMYD	GRAAGGVYAGDKEGDVIALMYD-----	1	1	0
2156	GDKEGDVTALMYDH	-----GDKEGDVTALMYDH-----	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
-----DKEGDVTALMYDHG-----	X33947	HSV-1-pUL19	1217	14aa	mRNA_30cpc_Organoid_MD114	1
-----DKEGDVTALMYDHG-----	X33947	HSV-2-pUL19	1217	14aa	mRNA_30cpc_Organoid_MD114	1
-----DKEGDVTALMYDHG-----	X33947	HSV-1-pUL19	1217	14aa	mRNA_3000cpc_Organoid_MD101	2
-----DKEGDVTALMYDHG-----	X33947	HSV-2-pUL19	1217	14aa	mRNA_3000cpc_Organoid_MD101	2
-----GDKEGDVTALMYDH-----	X57601	HSV-1-pUL19	1216	14aaA5	mRNA_3000cpc_Organoid_MD101	1
-----GDKEGDVTALMYDH-----	X57601	HSV-2-pUL19	1216	14aaA5	mRNA_3000cpc_Organoid_MD101	1
GRAAGGVYAGDKEGDVIALMYD-----	X62981	HSV-1-pUL19	1207	22aa	mRNA_30cpc_Organoid_MD114	1
GRAAGGVYAGDKEGDVIALMYD-----	X62981	HSV-2-pUL19	1207	22aa	mRNA_30cpc_Organoid_MD114	1
-----VIALMYDHGQSDPA-----	X66720	HSV-1-pUL19	1222	14aa	mRNA_30cpc_Organoid_MD114	1
-----VIALMYDHGQSDPA-----	X66720	HSV-2-pUL19	1222	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2147

Peptide AGADHVNAPNFPP from HSV-1-pUL19 with cluster number 2147

Peptide AGADHVNAPNFPP from HSV-1-pUL19 with cluster number 214



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2147	AGADHVNAPNFPP	6	2	4

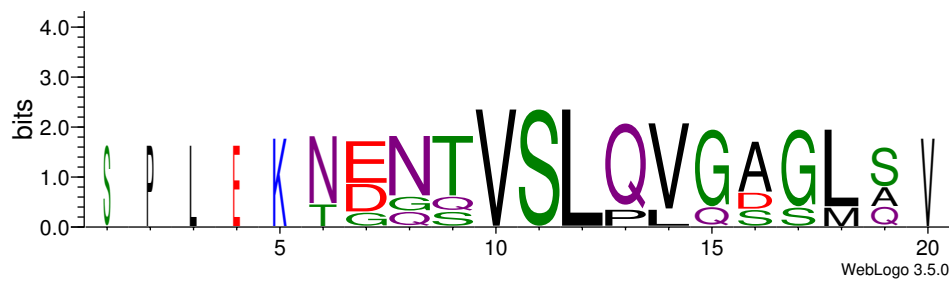
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2147	AGADHVNAPNFPP	_____AGADHVNAPNFPP_	2	0	2
2147	NEGADGVGNASGNW	_____NEGADGVGNASGNW__	1	1	0
2147	LPVHALFAGADHVNAPNFPPA	LPVHALFAGADHVNAPNFPPA	1	0	1
2147	GADHVNAPNFPPA	_____GADHVNAPNFPPA	1	0	1
2147	ALFAGADHVNAPN	_____ALFAGADHVNAPN_____	1	1	0

alignment	LUTnrs	GeneName	start	structure	Group	Score
_____AGADHVNAPNFPP_	X44203	HSV-1-pUL19	965	14aa	mRNA_3000cpc_Organoid_MD101	2
_____AGADHVNAPNFPP_	X44203	HSV-2-pUL19	965	14aa	mRNA_3000cpc_Organoid_MD101	2
_____ALFAGADHVNAPN_____	X42812	HSV-1-pUL19	962	14aa	mRNA_30cpc_Organoid_MD114	1
_____ALFAGADHVNAPN_____	X42812	HSV-2-pUL19	962	14aa	mRNA_30cpc_Organoid_MD114	1
_____GADHVNAPNFPPA	X64209	HSV-1-pUL19	966	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GADHVNAPNFPPA	X64209	HSV-2-pUL19	966	14aa	mRNA_3000cpc_Organoid_MD101	1
LPVHALFAGADHVNAPNFPPA	X30384	HSV-1-pUL19	958	22aa	mRNA_3000cpc_Organoid_MD101	1
LPVHALFAGADHVNAPNFPPA	X30384	HSV-2-pUL19	958	22aa	mRNA_3000cpc_Organoid_MD101	1
_____NEGADGVGNASGNW__	X2007	AAV1-VP1	215	14aa	mRNA_30cpc_Organoid_MD114	1
_____NEGADGVGNASGNW__	X2007	AAV6-VP1	215	14aa	mRNA_30cpc_Organoid_MD114	1
_____NEGADGVGNASGNW__	X2007	AAV8-VP1	216	14aa	mRNA_30cpc_Organoid_MD114	1
_____NEGADGVGNASGNW__	X2007	AAV9-VP1	215	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2003

Peptide NENTVSLQVGAGLS from CAV-2-F-SH01 with cluster number 2003

Peptide NENTVSLQVGAGLS from CAV-2-F-SH01 with cluster number



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2003	NENTVSLQVGAGLS	6	2	4

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2003	NENTVSLQVGAGLS	_____NENTVSLQVGAGLS_	2	1	1
2003	SPLEKNENTVSLQV	SPLEKNENTVSLQV_____	1	1	0
2003	KTDQQVSLQVGSG	____KTDQQVSLQVGSGL__	1	0	1
2003	GGTVSLPLQDSMQV	_____GGTVSLPLQDSMQV	1	0	1
2003	DNSVSLQVGAGLAV	_____DNSVSLQVGAGLAV	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
_____DNSVSLQVGAGLAV	X33725	CAV-2-F-SH01	218	14aa	mRNA_3000cpc_Organoid_MD101	1
_____DNSVSLQVGAGLAV	X33725	CAV-2-F-SH01	254	14aa	mRNA_3000cpc_Organoid_MD101	1
_____DNSVSLQVGAGLAV	X33725	CAV-2-F-Toronto-A26-61	218	14aa	mRNA_3000cpc_Organoid_MD101	1
_____DNSVSLQVGAGLAV	X33725	CAV-2-F-Toronto-A26-61	254	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GGTVSLPLQDSMQV	X64432	CAV-2-F-SH01	149	14aa	mRNA_3000cpc_Organoid_MD101	1
_____GGTVSLPLQDSMQV	X64432	CAV-2-F-Toronto-A26-61	149	14aa	mRNA_3000cpc_Organoid_MD101	1
____KTDQQVSLQVGSGL__	X3377	CAV-2-F-SH01	180	14aa	mRNA_3000cpc_Organoid_MD101	1
____KTDQQVSLQVGSGL__	X3377	CAV-2-F-Toronto-A26-61	180	14aa	mRNA_3000cpc_Organoid_MD101	1
_____NENTVSLQVGAGLS_	X1903	CAV-2-F-SH01	253	14aa	mRNA_30cpc_Organoid_MD114	1
_____NENTVSLQVGAGLS_	X1903	CAV-2-F-Toronto-A26-61	253	14aa	mRNA_30cpc_Organoid_MD114	1
_____NENTVSLQVGAGLS_	X1903	CAV-2-F-SH01	253	14aa	mRNA_3000cpc_Organoid_MD101	1
_____NENTVSLQVGAGLS_	X1903	CAV-2-F-Toronto-A26-61	253	14aa	mRNA_3000cpc_Organoid_MD101	1
SPLEKNENTVSLQV_____	X11647	CAV-2-F-SH01	248	14aa	mRNA_30cpc_Organoid_MD114	1
SPLEKNENTVSLQV_____	X11647	CAV-2-F-Toronto-A26-61	248	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2829

Peptide DEAGLQPKPLTTPP from HSV-1-pUL1 with cluster number 2829

Peptide DEAGLQPKPLTTPP from HSV-1-pUL1 with cluster number 2829



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2829	DEAGLQPKPLTTPP	5	1	4

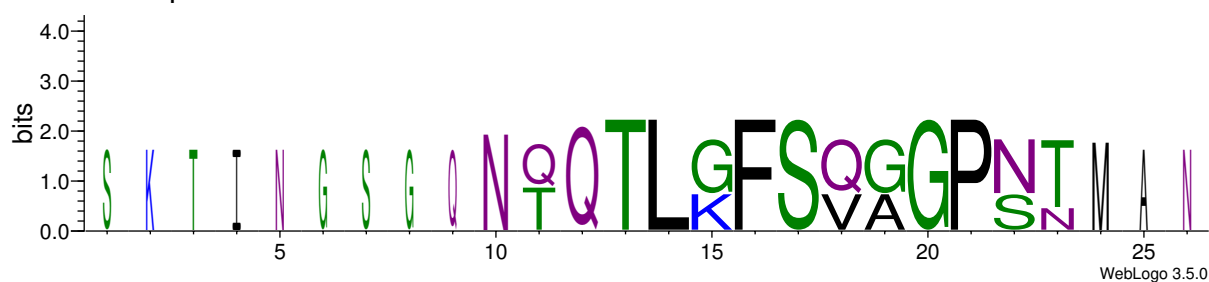
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2829	WEPPVSSDDEASSQSKPLATQP	WEPPVSSDDEASSQSKPLATQP	1	0	1
2829	SSDDEASSQSKPLA	____SSDDEASSQSKPLA____	1	0	1
2829	PVLPPDDEAGLQPK	__PVLPPDDEAGLQPK____	1	0	1
2829	PPDDEAGLQPKPLT	____PPDDEAGLQPKPLT____	1	0	1
2829	DEAGLQPKPLTTPP	____DEAGLQPKPLTTPP	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
____DEAGLQPKPLTTPP	X35725	HSV-1-pUL1	183	14aa	mRNA_30cpc_Organoid_MD114	1
____PPDDEAGLQPKPLT____	X22098	HSV-1-pUL1	180	14aa	mRNA_3000cpc_Organoid_MD101	1
__PVLPPDDEAGLQPK____	X24206	HSV-1-pUL1	177	14aa	mRNA_3000cpc_Organoid_MD101	1
____SSDDEASSQSKPLA____	X10780	HSV-2-pUL1	180	14aa	mRNA_3000cpc_Organoid_MD101	1
WEPPVSSDDEASSQSKPLATQP	X73978	HSV-2-pUL1	175	22aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2823

Peptide NQQLKFSVAGPSN from AAV9-VP1 with cluster number 2823

Peptide NQQLKFSVAGPSN from AAV9-VP1 with cluster number 2823



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2823	NQQLKFSVAGPSN	5	3	2

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2823	TQTLGFSQGGPNTM	-----TQTLGFSQGGPNTM--	1	1	0
2823	TLGFSQGGPNTMAN	-----TLGFSQGGPNTMAN	1	0	1
2823	SKTINGSGQNQQLKFSVAGPS	SKTINGSGQNQQLKFSVAGPS----	1	1	0
2823	NTQTLGFSQGGPNT	-----NTQTLGFSQGGPNT--	1	0	1
2823	NQQLKFSVAGPSN	-----NQQLKFSVAGPSN---	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
-----NQQLKFSVAGPSN---	X1142	AAV9-VP1	457	14aa	mRNA_30cpc_Organoid_MD114	1
-----NTQTLGFSQGGPNT---	X463	AAV8-VP1	459	14aa	mRNA_3000cpc_Organoid_MD101	1
SKTINGSGQNQQLKFSVAGPS----	X10032	AAV9-VP1	448	22aa	mRNA_30cpc_Organoid_MD114	1
-----TLGFSQGGPNTMAN	X7945	AAV8-VP1	462	14aa	mRNA_3000cpc_Organoid_MD101	1
-----TQTLGFSQGGPNTM--	X6974	AAV8-VP1	460	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2815

Peptide NFTLDARGDGAVVA from PRV-Becker-gE with cluster number 2815

» NFTLDARGDGAVVA from PRV-Becker-gE with cluster number



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2815	NFTLDARGDGAVVA	5	2	3

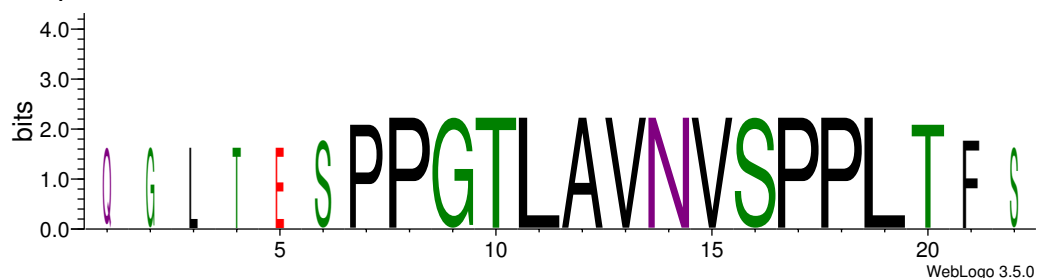
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2815	NFTLDARGDGAVVA	_NFTLDARGDGAVVA__	2	1	1
2815	TLDARGDGAVVAGI	___TLDARGDGAVVAGI	1	0	1
2815	ANFTLDARGDGAVV	ANFTLDARGDGAVV___	1	1	0
2815	ANFTLDARGDGAVL	ANFTLDARGDGAVL___	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
ANFTLDARGDGAVL___	X39958	PRV-Becker-gE	92	14aa	mRNA_3000cpc_Organoid_MD101	1
ANFTLDARGDGAVL___	X39958	PRV-TJ-gE	93	14aa	mRNA_3000cpc_Organoid_MD101	1
ANFTLDARGDGAVV___	X39959	PRV-Becker-gE	92	14aa	mRNA_30cpc_Organoid_MD114	1
ANFTLDARGDGAVV___	X39959	PRV-TJ-gE	93	14aa	mRNA_30cpc_Organoid_MD114	1
_NFTLDARGDGAVVA__	X2904	PRV-Becker-gE	93	14aa	mRNA_30cpc_Organoid_MD114	1
_NFTLDARGDGAVVA__	X2904	PRV-TJ-gE	94	14aa	mRNA_30cpc_Organoid_MD114	1
_NFTLDARGDGAVVA__	X2904	PRV-Becker-gE	93	14aa	mRNA_3000cpc_Organoid_MD101	1
_NFTLDARGDGAVVA__	X2904	PRV-TJ-gE	94	14aa	mRNA_3000cpc_Organoid_MD101	1
___TLDARGDGAVVAGI	X7877	PRV-Becker-gE	95	14aa	mRNA_3000cpc_Organoid_MD101	1
___TLDARGDGAVVAGI	X7877	PRV-TJ-gE	96	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2727

Peptide PPGTLAVNVSPPLT from CAV-2-F-SH01 with cluster number 2727

Peptide PPGTLAVNVSPPLT from CAV-2-F-SH01 with cluster number 2727



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2727	PPGTLAVNVSPPLT	5	1	4

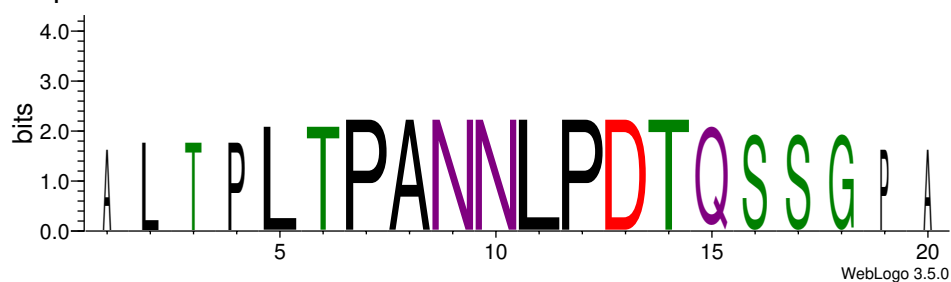
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2727	PPGTLAVNVSPPLT	_____PPGTLAVNVSPPLT__	2	1	1
2727	SPPGTLAVNVSPPL	_____SPPGTLAVNVSPPL__	1	0	1
2727	QGLTESPPGTLAVNVSPPLTFS	QGLTESPPGTLAVNVSPPLTFS	1	0	1
2727	PGTLAVNVSPPLTF	_____PGTLAVNVSPPLTF_	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
_____PGTLAVNVSPPLTF_	X23756	CAV-2-F-SH01	41	14aa	mRNA_3000cpc_Organoid_MD101	1
_____PGTLAVNVSPPLTF_	X23756	CAV-2-F-Toronto-A26-61	41	14aa	mRNA_3000cpc_Organoid_MD101	1
_____PPGTLAVNVSPPLT__	X22192	CAV-2-F-SH01	40	14aa	mRNA_30cpc_Organoid_MD114	1
_____PPGTLAVNVSPPLT__	X22192	CAV-2-F-Toronto-A26-61	40	14aa	mRNA_30cpc_Organoid_MD114	1
_____PPGTLAVNVSPPLT__	X22192	CAV-2-F-SH01	40	14aa	mRNA_3000cpc_Organoid_MD101	1
_____PPGTLAVNVSPPLT__	X22192	CAV-2-F-Toronto-A26-61	40	14aa	mRNA_3000cpc_Organoid_MD101	1
QGLTESPPGTLAVNVSPPLTFS	X20051	CAV-2-F-SH01	34	22aa	mRNA_3000cpc_Organoid_MD101	1
QGLTESPPGTLAVNVSPPLTFS	X20051	CAV-2-F-Toronto-A26-61	34	22aa	mRNA_3000cpc_Organoid_MD101	1
_____SPPGTLAVNVSPPL__	X11599	CAV-2-F-SH01	39	14aa	mRNA_3000cpc_Organoid_MD101	1
_____SPPGTLAVNVSPPL__	X11599	CAV-2-F-Toronto-A26-61	39	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2626

Peptide LTPANNLPDTQSSG from PV2-VP1 with cluster number 2626

Peptide LTPANNLPDTQSSG from PV2-VP1 with cluster number 2626



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2626	LTPANNLPDTQSSG	5	0	5

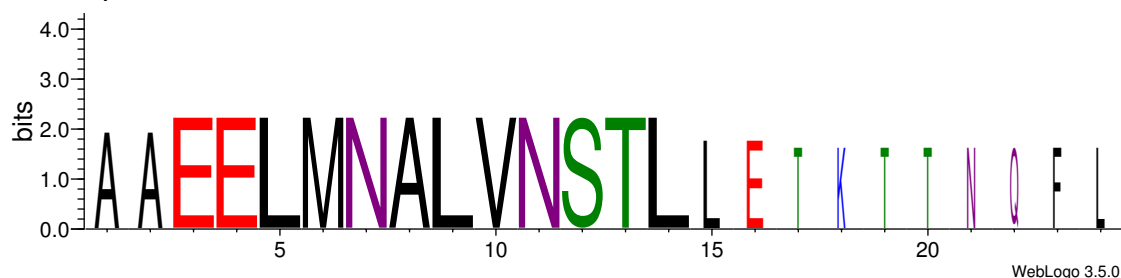
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2626	LTPANNLPDTQSSG	___LTPANNLPDTQSSG__	2	0	2
2626	PANNLPDTQSSGPA	____PANNLPDTQSSGPA	1	0	1
2626	LTPLTPANNLPDTQ	_LTPLTPANNLPDTQ_____	1	0	1
2626	ALTPLTPANNLPDT	ALTPLTPANNLPDT_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
ALTPLTPANNLPDT_____	X42333	PV2-VP1	17	14aa	mRNA_3000cpc_Organoid_MD101	1
___LTPANNLPDTQSSG__	X28917	PV2-VP1	21	14aa	mRNA_3000cpc_Organoid_MD101	2
_LTPLTPANNLPDTQ_____	X28902	PV2-VP1	18	14aa	mRNA_3000cpc_Organoid_MD101	1
____PANNLPDTQSSGPA	X23293	PV2-VP1	23	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2528

Peptide AAEELMNALVNSTL from MV-H-CAM-70 with cluster number 2528

Peptide AAEELMNALVNSTL from MV-H-CAM-70 with cluster number 2528



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2528	AAEELMNALVNSTL	5	2	3

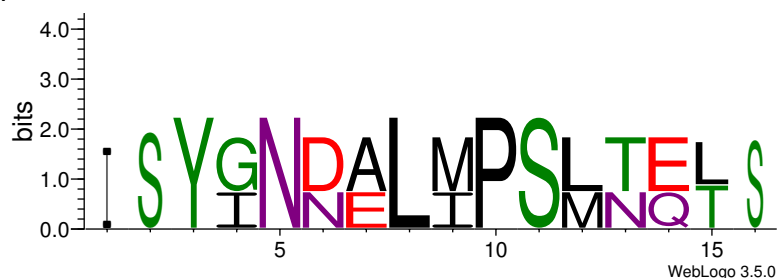
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2528	AAEELMNALVNSTL	AAEELMNALVNSTL_____	3	1	2
2528	EELMNALVNSTLLETKTNNQFL	__EELMNALVNSTLLETKTNNQFL	1	1	0
2528	EELMNALVNSTLLE	__EELMNALVNSTLLE_____	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
AAEELMNALVNSTL_____	X43711	MV-H-CAM-70	158	14aa	mRNA_30cpc_Organoid_MD114	1
AAEELMNALVNSTL_____	X43711	MV-H-Edmonston	158	14aa	mRNA_30cpc_Organoid_MD114	1
AAEELMNALVNSTL_____	X43711	MV-H-Halle	158	14aa	mRNA_30cpc_Organoid_MD114	1
AAEELMNALVNSTL_____	X43711	MV-H-MVi-Illinois.USA-50.99	158	14aa	mRNA_30cpc_Organoid_MD114	1
AAEELMNALVNSTL_____	X43711	MV-H-MVi-Texas.USA-28.99	158	14aa	mRNA_30cpc_Organoid_MD114	1
AAEELMNALVNSTL_____	X43711	MV-H-CAM-70	158	14aa	mRNA_3000cpc_Organoid_MD101	2
AAEELMNALVNSTL_____	X43711	MV-H-Edmonston	158	14aa	mRNA_3000cpc_Organoid_MD101	2
AAEELMNALVNSTL_____	X43711	MV-H-Halle	158	14aa	mRNA_3000cpc_Organoid_MD101	2
AAEELMNALVNSTL_____	X43711	MV-H-MVi-Illinois.USA-50.99	158	14aa	mRNA_3000cpc_Organoid_MD101	2
AAEELMNALVNSTL_____	X43711	MV-H-MVi-Texas.USA-28.99	158	14aa	mRNA_3000cpc_Organoid_MD101	2
__EELMNALVNSTLLE_____	X38580	MV-H-CAM-70	160	14aa	mRNA_3000cpc_Organoid_MD101	1
__EELMNALVNSTLLE_____	X38580	MV-H-Edmonston	160	14aa	mRNA_3000cpc_Organoid_MD101	1
__EELMNALVNSTLLE_____	X38580	MV-H-Halle	160	14aa	mRNA_3000cpc_Organoid_MD101	1
__EELMNALVNSTLLE_____	X38580	MV-H-MVi-Illinois.USA-50.99	160	14aa	mRNA_3000cpc_Organoid_MD101	1
__EELMNALVNSTLLE_____	X38580	MV-H-MVi-Texas.USA-28.99	160	14aa	mRNA_3000cpc_Organoid_MD101	1
__EELMNALVNSTLLETKTNNQFL	X38578	MV-H-CAM-70	160	22aa	mRNA_30cpc_Organoid_MD114	1
__EELMNALVNSTLLETKTNNQFL	X38578	MV-H-Edmonston	160	22aa	mRNA_30cpc_Organoid_MD114	1
__EELMNALVNSTLLETKTNNQFL	X38578	MV-H-Halle	160	22aa	mRNA_30cpc_Organoid_MD114	1
__EELMNALVNSTLLETKTNNQFL	X38578	MV-H-MVi-Illinois.USA-50.99	160	22aa	mRNA_30cpc_Organoid_MD114	1
__EELMNALVNSTLLETKTNNQFL	X38578	MV-H-MVi-Texas.USA-28.99	160	22aa	mRNA_30cpc_Organoid_MD114	1

[1] 2524

Peptide SYGNDALMPSLTET from APP with cluster number 2524

ptide SYGNDALMPSLTET from APP with cluster number



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2524	SYGNDALMPSLTET	5	3	2

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2524	YINNELIPSMNQLS	__YINNELIPSMNQLS	2	2	0
2524	SYGNDALMPSLTET	_SYGNDALMPSLTET_	2	1	1
2524	ISYGNDALMPSLTE	ISYGNDALMPSLTE__	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
ISYGNDALMPSLTE__	X14903	APP	586	14aa	mRNA_3000cpc_Organoid_MD101	1
SYGNDALMPSLTET	X13829	APP	587	14aa	mRNA_30cpc_Organoid_MD114	1
SYGNDALMPSLTET	X13829	APP	587	14aa	mRNA_3000cpc_Organoid_MD101	1
__YINNELIPSMNQLS	X59499	MV-F-CAM-70	184	14aaA5	mRNA_30cpc_Organoid_MD114	1
__YINNELIPSMNQLS	X90789	MV-F-CAM-70	184	14aaG4S	mRNA_30cpc_Organoid_MD114	1
__YINNELIPSMNQLS	X59499	MV-F-Edmonston	184	14aaA5	mRNA_30cpc_Organoid_MD114	1
__YINNELIPSMNQLS	X90789	MV-F-Edmonston	184	14aaG4S	mRNA_30cpc_Organoid_MD114	1
__YINNELIPSMNQLS	X59499	MV-F-Halle	184	14aaA5	mRNA_30cpc_Organoid_MD114	1
__YINNELIPSMNQLS	X90789	MV-F-Halle	184	14aaG4S	mRNA_30cpc_Organoid_MD114	1

[1] 2490

Peptide ANSKVLIELEPPFG from WNV-E-ArD76104 with cluster number 2490

Peptide ANSKVLIELEPPFG from WNV-E-ArD76104 with cluster number 2490



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2490	ANSKVLIELEPPFG	5	5	0

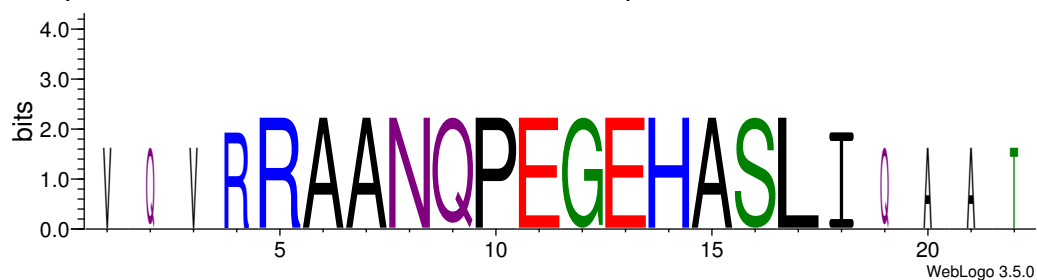
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2490	TANSKVLIELEPPF	TANSKVLIELEPPF_____	1	1	0
2490	SSNSKVLVGMEPPF	SSNSKVLVGMEPPF_____	1	1	0
2490	SKVLVEMEPPFGDS	__SKVLVEMEPPFGDS_____	1	1	0
2490	NSKVLIELEPPFGDSYIVVGRG	__NSKVLIELEPPFGDSYIVVGRG	1	1	0
2490	ANSKVLIELEPPFG	_ANSKVLIELEPPFG_____	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
_ANSKVLIELEPPFG_____	X39811	WNV-E-ArD76104	363	14aa	mRNA_30cpc_Organoid_MD114	1
_ANSKVLIELEPPFG_____	X39811	WNV-E-H-442	367	14aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	JEV-E-BN19	367	22aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	JEV-E-KE-093	367	22aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	JEV-E-SA14-5-3	367	22aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	JEV-E-ThCMAr4492	367	22aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	WNV-E-AZ10-918	368	22aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	WNV-E-ArD76104	364	22aa	mRNA_30cpc_Organoid_MD114	1
__NSKVLIELEPPFGDSYIVVGRG	X594	WNV-E-H-442	368	22aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	JEV-E-BN19	368	14aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	JEV-E-KE-093	368	14aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	JEV-E-SA14-5-3	368	14aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	JEV-E-ThCMAr4492	368	14aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	WNV-E-AZ10-918	369	14aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	WNV-E-ArD76104	365	14aa	mRNA_30cpc_Organoid_MD114	1
___SKVLVEMEPPFGDS_____	X10181	WNV-E-H-442	369	14aa	mRNA_30cpc_Organoid_MD114	1
SSNSKVLVGMEPPF_____	X10547	JEV-E-KE-093	365	14aa	mRNA_30cpc_Organoid_MD114	1
SSNSKVLVGMEPPF_____	X10547	JEV-E-ThCMAr4492	365	14aa	mRNA_30cpc_Organoid_MD114	1
TANSKVLIELEPPF_____	X8387	WNV-E-ArD76104	362	14aa	mRNA_30cpc_Organoid_MD114	1
TANSKVLIELEPPF_____	X8387	WNV-E-H-442	366	14aa	mRNA_30cpc_Organoid_MD114	1

[1] 2258

Peptide RAANQPEGEHASLI from HSV-1-pUL36 with cluster number 2258

Peptide RAANQPEGEHASLI from HSV-1-pUL36 with cluster number 225



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2258	RAANQPEGEHASLI	5	2	3

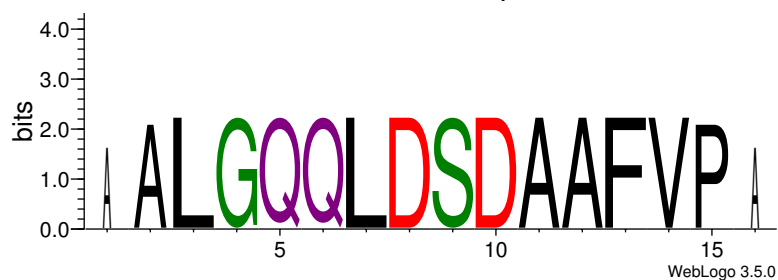
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2258	RRAANQPEGEHASL	___RRAANQPEGEHASL_____	2	0	2
2258	RAANQPEGEHASLI	___RAANQPEGEHASLI_____	2	1	1
2258	VQVRRRAANQPEGEHASLIQAAT	VQVRRRAANQPEGEHASLIQAAT	1	1	0

alignment	LUTnrs	GeneName	start	structure	Group	Score
___RAANQPEGEHASLI_____	X27243	HSV-1-pUL36	1235	14aa	mRNA_30cpc_Organoid_MD114	1
___RAANQPEGEHASLI_____	X27243	HSV-1-pUL36	1235	14aa	mRNA_3000cpc_Organoid_MD101	1
___RRAANQPEGEHASL_____	X26212	HSV-1-pUL36	1234	14aa	mRNA_3000cpc_Organoid_MD101	2
VQVRRRAANQPEGEHASLIQAAT	X67035	HSV-1-pUL36	1231	22aa	mRNA_30cpc_Organoid_MD114	1

[1] 2250

Peptide ALGQQLDSDAAFVP from HSV-1-pUL36 with cluster number 2250

ALGQQLDSDAAFVP from HSV-1-pUL36 with cluster nu



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2250	ALGQQLDSDAAFVP	5	3	2

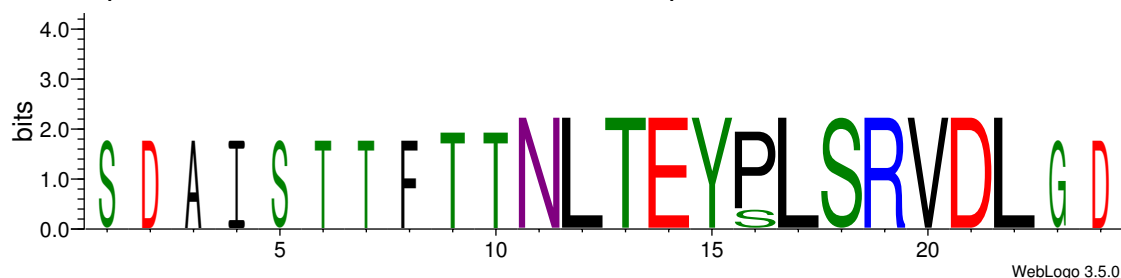
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2250	ALGQQLDSDAAFVP	_ALGQQLDSDAAFVP_	3	2	1
2250	LGQQLDSDAAFVPA	__LGQQLDSDAAFVPA	1	0	1
2250	AALGQQLDSDAAFV	AALGQQLDSDAAFV__	1	1	0

alignment	LUTnrs	GeneName	start	structure	Group	Score
AALGQQLDSDAAFV__	X43643	HSV-1-pUL36	933	14aa	mRNA_30cpc_Organoid_MD114	1
ALGQQLDSDAAFVP	X42683	HSV-1-pUL36	934	14aa	mRNA_30cpc_Organoid_MD114	1
ALGQQLDSDAAFVP	X56327	HSV-1-pUL36	934	14aaA5	mRNA_30cpc_Organoid_MD114	1
ALGQQLDSDAAFVP	X56327	HSV-1-pUL36	934	14aaA5	mRNA_3000cpc_Organoid_MD101	1
__LGQQLDSDAAFVPA	X32402	HSV-1-pUL36	935	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2219

Peptide NLTEYPLSRVDLGD from HSV-1-pUL27 with cluster number 2219

Peptide NLTEYPLSRVDLGD from HSV-1-pUL27 with cluster number 2219



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2219	NLTEYPLSRVDLGD	5	2	3

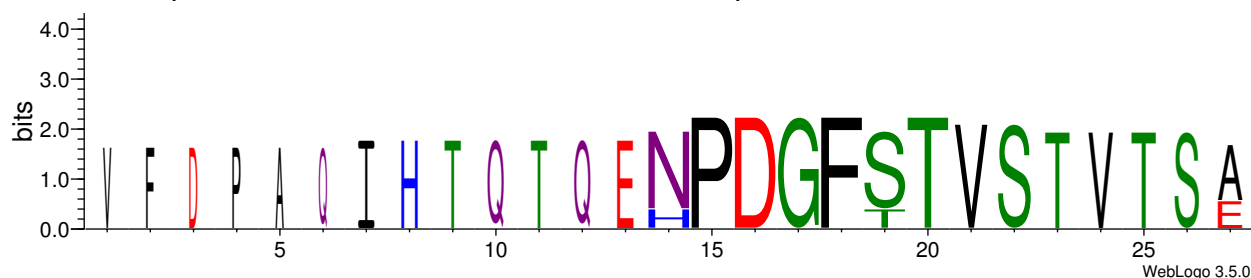
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2219	SDAISTTFTTNLTEYPLSRVDL	SDAISTTFTTNLTEYPLSRVDL__	2	1	1
2219	NLTEYPLSRVDLGD	_____NLTEYPLSRVDLGD	2	1	1
2219	TTNLTEYSLSRVDL	_____TTNLTEYSLSRVDL__	1	0	1

alignment	LUTnrs	GeneName	start	structure	Group	Score
_____NLTEYPLSRVDLGD	X1575	HSV-1-pUL27	398	14aa	mRNA_30cpc_Organoid_MD114	1
_____NLTEYPLSRVDLGD	X1575	HSV-2-pUL27	393	14aa	mRNA_30cpc_Organoid_MD114	1
_____NLTEYPLSRVDLGD	X1575	HSV-1-pUL27	398	14aa	mRNA_3000cpc_Organoid_MD101	1
_____NLTEYPLSRVDLGD	X1575	HSV-2-pUL27	393	14aa	mRNA_3000cpc_Organoid_MD101	1
SDAISTTFTTNLTEYPLSRVDL__	X12537	HSV-1-pUL27	388	22aa	mRNA_30cpc_Organoid_MD114	1
SDAISTTFTTNLTEYPLSRVDL__	X12537	HSV-2-pUL27	383	22aa	mRNA_30cpc_Organoid_MD114	1
SDAISTTFTTNLTEYPLSRVDL__	X12537	HSV-1-pUL27	388	22aa	mRNA_3000cpc_Organoid_MD101	1
SDAISTTFTTNLTEYPLSRVDL__	X12537	HSV-2-pUL27	383	22aa	mRNA_3000cpc_Organoid_MD101	1
_____TTNLTEYSLSRVDL__	X5884	HSV-1-pUL27	396	14aa	mRNA_3000cpc_Organoid_MD101	1
_____TTNLTEYSLSRVDL__	X5884	HSV-2-pUL27	391	14aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2056

Peptide NPDGFSTVSTVTSA from HSV-1-pUL44 with cluster number 2056

Peptide NPDGFSTVSTVTSA from HSV-1-pUL44 with cluster number 2056



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2056	NPDGFSTVSTVTSA	5	1	4

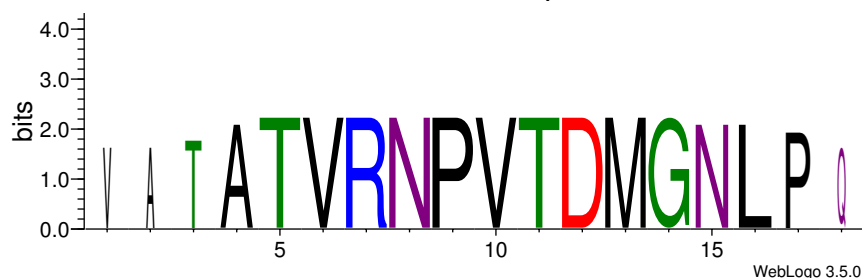
cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2056	NPDGFSTVSTVTSA	-----NPDGFSTVSTVTSA	2	1	1
2056	VFDPAQIHTQTQENPDGFSTVS	VFDPAQIHTQTQENPDGFSTVS----	1	0	1
2056	IHTQTQENPDGFST	-----IHTQTQENPDGFST-----	1	0	1
2056	HPDGFTTVSTVTSE	-----HPDGFTTVSTVTSE	1	0	1

alignment	LUThrs	GeneName	start	structure	Group	Score
-----HPDGFTTVSTVTSE	X17560	HSV-1-pUL44	321	14aa	mRNA_3000cpc_Organoid_MD101	1
-----HPDGFTTVSTVTSE	X17560	HSV-2-pUL44	290	14aa	mRNA_3000cpc_Organoid_MD101	1
-----IHTQTQENPDGFST-----	X15159	HSV-2-pUL44	283	14aa	mRNA_3000cpc_Organoid_MD101	1
-----NPDGFSTVSTVTSA	X1317	HSV-1-pUL44	321	14aa	mRNA_30cpc_Organoid_MD114	1
-----NPDGFSTVSTVTSA	X1317	HSV-2-pUL44	290	14aa	mRNA_30cpc_Organoid_MD114	1
-----NPDGFSTVSTVTSA	X1317	HSV-1-pUL44	321	14aa	mRNA_3000cpc_Organoid_MD101	1
-----NPDGFSTVSTVTSA	X1317	HSV-2-pUL44	290	14aa	mRNA_3000cpc_Organoid_MD101	1
VFDPAQIHTQTQENPDGFSTVS----	X69975	HSV-2-pUL44	277	22aa	mRNA_3000cpc_Organoid_MD101	1

[1] 2021

Peptide ATVRNPVTDMGNLP from HSV-1-pUL19 with cluster number 2021

ide ATVRNPVTDMGNLP from HSV-1-pUL19 with cluster numb



cluster_id	main_sequence	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2021	ATVRNPVTDMGNLP	5	4	1

cluster_id	sequence	alignment	sum	mRNA_30cpc_Organoid_MD114	mRNA_3000cpc_Organoid_MD101
2021	ATVRNPVTDMGNLP	___ATVRNPVTDMGNLP_	2	1	1
2021	VATATVRNPVTDMG	VATATVRNPVTDMG_____	1	1	0
2021	TVRNPVTDMGNLPQ	_____TVRNPVTDMGNLPQ	1	1	0
2021	TATVRNPVTDMGNL	__TATVRNPVTDMGNL__	1	1	0

alignment	LUThrs	GeneName	start	structure	Group	Score
___ATVRNPVTDMGNLP_	X40482	HSV-1-pUL19	1113	14aa	mRNA_30cpc_Organoid_MD114	1
___ATVRNPVTDMGNLP_	X40482	HSV-2-pUL19	1113	14aa	mRNA_30cpc_Organoid_MD114	1
___ATVRNPVTDMGNLP_	X40482	HSV-1-pUL19	1113	14aa	mRNA_3000cpc_Organoid_MD101	1
___ATVRNPVTDMGNLP_	X40482	HSV-2-pUL19	1113	14aa	mRNA_3000cpc_Organoid_MD101	1
__TATVRNPVTDMGNL__	X8426	HSV-1-pUL19	1112	14aa	mRNA_30cpc_Organoid_MD114	1
__TATVRNPVTDMGNL__	X8426	HSV-2-pUL19	1112	14aa	mRNA_30cpc_Organoid_MD114	1
_____TVRNPVTDMGNLPQ	X9157	HSV-1-pUL19	1114	14aa	mRNA_30cpc_Organoid_MD114	1
_____TVRNPVTDMGNLPQ	X9157	HSV-2-pUL19	1114	14aa	mRNA_30cpc_Organoid_MD114	1
VATATVRNPVTDMG_____	X68514	HSV-1-pUL19	1110	14aa	mRNA_30cpc_Organoid_MD114	1
VATATVRNPVTDMG_____	X68514	HSV-2-pUL19	1110	14aa	mRNA_30cpc_Organoid_MD114	1

```
# setkey(select.samples.trsp,Peptide) select.samples.trsp.select <-
# select.samples.trsp[J(c('PPDELNLTTASLPL'))]
```

```
print("Total analysis time:")
```

```
[1] "Total analysis time:"
```

```
print(Sys.time() - strt1)
```

```
Time difference of 1.529664 mins
```

```
devtools::session_info()
```

```
Session info -----
```

```
setting  value
version  R version 3.4.2 (2017-09-28)
system   x86_64, linux-gnu
ui        X11
language (EN)
collate   en_US.UTF-8
tz        UTC
date      2020-11-09
```

```
Packages -----
```

package	* version	date	source
ade4	1.7-8	2017-08-09	CRAN (R 3.4.2)
backports	1.1.1	2017-09-25	CRAN (R 3.4.2)
base	* 3.4.2	2017-10-06	local
compiler	3.4.2	2017-10-06	local
data.table	* 1.10.4-2	2017-10-12	url
datasets	* 3.4.2	2017-10-06	local
devtools	* 1.13.3	2017-08-02	CRAN (R 3.4.2)
digest	0.6.12	2017-01-27	CRAN (R 3.4.2)
evaluate	0.10.1	2017-06-24	CRAN (R 3.4.2)
formatR	1.5	2017-04-25	CRAN (R 3.4.2)
graphics	* 3.4.2	2017-10-06	local
grDevices	* 3.4.2	2017-10-06	local
highr	0.6	2016-05-09	CRAN (R 3.4.2)
hms	0.3	2016-11-22	CRAN (R 3.4.2)
htmltools	0.3.6	2017-04-28	CRAN (R 3.4.2)
httr	1.3.1	2017-08-20	CRAN (R 3.4.2)
kableExtra	* 0.5.2	2017-09-15	url
knitr	* 1.17	2017-08-10	CRAN (R 3.4.2)
magrittr	1.5	2014-11-22	CRAN (R 3.4.2)
memoise	1.1.0	2017-04-21	CRAN (R 3.4.2)
methods	3.4.2	2017-10-06	local
parallel	* 3.4.2	2017-10-06	local
plyr	1.8.4	2016-06-08	CRAN (R 3.4.2)
R6	2.2.2	2017-06-17	CRAN (R 3.4.2)
Rcpp	0.12.13	2017-09-28	url
readr	1.1.1	2017-05-16	CRAN (R 3.4.2)
reshape2	* 1.4.2	2016-10-22	CRAN (R 3.4.2)
rlang	0.1.2	2017-08-09	CRAN (R 3.4.2)
rmarkdown	1.6	2017-06-15	url
rprojroot	1.2	2017-01-16	CRAN (R 3.4.2)
rvest	0.3.2	2016-06-17	CRAN (R 3.4.2)
seqinr	* 3.4-5	2017-08-01	CRAN (R 3.4.2)

stats	* 3.4.2	2017-10-06	local
stringi	1.1.5	2017-04-07	url
stringr	1.2.0	2017-02-18	CRAN (R 3.4.2)
tibble	1.3.4	2017-08-22	CRAN (R 3.4.2)
tools	3.4.2	2017-10-06	local
utils	* 3.4.2	2017-10-06	local
withr	2.0.0	2017-07-28	url
xml2	1.1.1	2017-01-24	CRAN (R 3.4.2)
yaml	2.1.14	2016-11-12	CRAN (R 3.4.2)