```
> if (!require("BiocManager", quietly = TRUE))
    install.packages("BiocManager")
> BiocManager::install("Biostrings")

> library(ape)
> library(Biostrings)
> library(seqinr)

> Papio_hamadryas.fas =
read.dna("~/MATERIALE_DIDATTICO_A.S._2024-2025/ESERCITAZIONI/MIOSTATINA/
Papio_hamadryas_myostatin.fas", format="fasta")

> Papio_hamadryas.fas
1 DNA sequence in binary format stored in a matrix.

Sequence length: 1128

Label:
AF019619.1 Papio hamadryas myostatin (MSTN) mRNA, complete c...

Base composition:
    a     c     g     t
0.330 0.199 0.210 0.262
(Total: 1.13 kb)

> str(Papio_hamadryas.fas)
 'DNAbin' raw [1, 1:1128] a t g c ...
 - attr(*, "dimnames")=List of 2
  ..$ : chr "AF019619.1 Papio hamadryas myostatin (MSTN) mRNA, complete cds"
  ..$ : NULL

> Papio_hamadryas.fas[1]
1 DNA sequence in binary format stored in a vector.

Sequence length: 1

Base composition:
a c g t
1 0 0 0

> Papio_hamadryas.fas[seq(3, length(Papio_hamadryas.fas), by = 3)]
1 DNA sequence in binary format stored in a vector.

Sequence length: 376

Base composition:
    a     c     g     t
0.324 0.181 0.170 0.324
(Total: 376 bases)


> base.freq(Papio_hamadryas.fas)
        a         c         g         t
0.3297872 0.1985816 0.2101064 0.2615248


> comp(Papio_hamadryas.fas)
  [1] "t" "a" "c" "g" "t" "t" "t" "t" "t" "g" "a" "c" "g" "t" "t" "g" "a" "g"
 [19] "a" "c" "a" "c" "a" "a" "a" "t" "a" "t" "a" "a" "a" "t" "g" "g" "a" "c"
 [37] "a" "a" "a" "t" "a" "c" "g" "a" "c" "t" "a" "a" "c" "a" "a" "c" "g" "a"
 ...

> c2s(Papio_hamadryas.fas)
```

```
[1]
"atgcaaaaactgcaactctgtgtttatatttacctgtttatgctgattgttgctggtccagtggatctaaatgagaacagtgagcaaa
aagaaaatgtggaaaaagaggggctgtgtaatgcatgtacttggagacaaaacactaaatcttcaagaatagaagccattaaaatacaa
atcctcagtaaacttcgtctggaaacagctcctaacatcagcaaagatgctataagacaacttttacccaaagctcctccactccggga
actgattgatcagtatgatgtccagagggatgacagcagcgatggctctttggaagatgacgattatcacgctacaacggaaacaatca
ttaccatgcctacagagtctgattttttaatgcaagtggatggaaaacccaaatgttgcttctttaaatttagctctaaaatacaatac
aataaagtggtaaaggcccaactatggatatatttgagacccgtcgagactcctacaacagtgtttgtgcaaatcctgagactcatcaa
acctatgaaagacggtacaaggtatactggaatccgatctctgaaacttgacatgaacccaggcactggtatttggcagagcattgatg
tgaagacagtgttgcaaaattggctcaaacaacctgaatccaacttaggcattgaaataaaagctttagatgagaatggtcatgatctt
gctgtaaccttcccaggaccaggagaagatgggctgaatccctttttagaggtcaaggtaacagacacaccaaaaagatccagaaggga
ttttggtcttgactgtgatgagcactcaacagaatcgcgatgctgtcgttaccctctaactgtggattttgaagctcttggatgggatt
ggattatcgctcctaaaagatataaggccaattactgctctggagagtgtgaatttgtatttttacaaaaatatcctcatactcatctg
gtacaccaagcaaaccccagaggttcagcaggcccttgctgtactcccacaaagatgtctccaattaatatgctatattttaatggcaa
agaacaaataatatatgggaaaattccagccatggtagtagaccgctgcgggtgctcatga"
```

> **Papio_hamadryas_string = c2s(Papio_hamadryas.fas)**

> **Papio_hamadryas_string**
```
[1]
"atgcaaaaactgcaactctgtgtttatatttacctgtttatgctgattgttgctggtccagtggatctaaatgagaacagtgagcaaa
aagaaaatgtggaaaaagaggggctgtgtaatgcatgtacttggagacaaaacactaaatcttcaagaatagaagccattaaaatacaa
atcctcagtaaacttcgtctggaaacagctcctaacatcagcaaagatgctataagacaacttttacccaaagctcctccactccggga
actgattgatcagtatgatgtccagagggatgacagcagcgatggctctttggaagatgacgattatcacgctacaacggaaacaatca
ttaccatgcctacagagtctgattttttaatgcaagtggatggaaaacccaaatgttgcttctttaaatttagctctaaaatacaatac
aataaagtggtaaaggcccaactatggatatatttgagacccgtcgagactcctacaacagtgtttgtgcaaatcctgagactcatcaa
acctatgaaagacggtacaaggtatactggaatccgatctctgaaacttgacatgaacccaggcactggtatttggcagagcattgatg
tgaagacagtgttgcaaaattggctcaaacaacctgaatccaacttaggcattgaaataaaagctttagatgagaatggtcatgatctt
gctgtaaccttcccaggaccaggagaagatgggctgaatccctttttagaggtcaaggtaacagacacaccaaaaagatccagaaggga
ttttggtcttgactgtgatgagcactcaacagaatcgcgatgctgtcgttaccctctaactgtggattttgaagctcttggatgggatt
ggattatcgctcctaaaagatataaggccaattactgctctggagagtgtgaatttgtatttttacaaaaatatcctcatactcatctg
gtacaccaagcaaaccccagaggttcagcaggcccttgctgtactcccacaaagatgtctccaattaatatgctatattttaatggcaa
agaacaaataatatatgggaaaattccagccatggtagtagaccgctgcgggtgctcatga"
```

> **str(Papio_hamadryas_string)**
```
 chr
"atgcaaaaactgcaactctgtgtttatatttacctgtttatgctgattgttgctggtccagtggatctaaatgagaacagtgagcaaa
aagaaaatgtggaaaaagagggg"| __truncated__
```

> **length(Papio_hamadryas_string)**
```
[1] 1
```

> **Papio_hamadryas_vector = s2c(Papio_hamadryas_string)**

> **Papio_hamadryas_vector**
```
  [1] "a" "t" "g" "c" "a" "a" "a" "a" "a" "c" "t" "g" "c" "a" "a" "c" "t" "c"
 [19] "t" "g" "t" "g" "t" "t" "t" "a" "t" "a" "t" "t" "t" "a" "c" "c" "t" "g"
 [37] "t" "t" "t" "a" "t" "g" "c" "t" "g" "a" "t" "t" "g" "t" "t" "g" "c" "t"
 ...
```

> **str(Papio_hamadryas_vector)**
```
 chr [1:1128] "a" "t" "g" "c" "a" "a" "a" "a" "a" "c" "t" "g" "c" "a" "a" ...
```

> **length(Papio_hamadryas_vector)**
```
[1] 1128
```

> **library(seqinr)**

> **rev(Papio_hamadryas_vector)**
```
  [1] "a" "g" "t" "a" "c" "t" "c" "g" "t" "g" "g" "g" "c" "g" "t" "c" "g" "c"
 [19] "c" "a" "g" "a" "t" "g" "a" "t" "g" "g" "t" "a" "c" "c" "g" "a" "c" "c"
 [37] "t" "t" "a" "a" "a" "a" "g" "g" "g" "t" "a" "t" "a" "t" "a" "a" "t" "a"
 ...
```

> **comp(Papio_hamadryas_vector)**

```
   [1] "t" "a" "c" "g" "t" "t" "t" "t" "t" "g" "a" "c" "g" "t" "t" "g" "a" "g"
  [19] "a" "c" "a" "c" "a" "a" "a" "t" "a" "t" "a" "a" "a" "t" "g" "g" "a" "c"
  [37] "a" "a" "a" "t" "a" "c" "g" "a" "c" "t" "a" "a" "c" "a" "a" "c" "g" "a"
  [55] "c" "c" "a" "g" "g" "t" "c" "a" "c" "c" "t" "a" "g" "a" "t" "t" "t" "a"
  ...

> Papio_hamadryas_vector
   [1] "a" "t" "g" "c" "a" "a" "a" "a" "a" "c" "t" "g" "c" "a" "a" "c" "t" "c"
  [19] "t" "g" "t" "g" "t" "t" "t" "a" "t" "a" "t" "t" "t" "a" "c" "c" "t" "g"
  [37] "t" "t" "t" "a" "t" "g" "c" "t" "g" "a" "t" "t" "g" "t" "t" "g" "c" "t"
  ...

> Papio_hamadryas_vector[seq(3, length(Papio_hamadryas_vector), by = 3)]
   [1] "g" "a" "a" "g" "a" "c" "t" "t" "t" "t" "c" "g" "t" "g" "g" "t" "t" "t"
  [19] "t" "a" "g" "t" "a" "t" "g" "c" "t" "g" "a" "a" "a" "t" "g" "a" "a" "g"
  [37] "g" "g" "t" "t" "a" "t" "t" "g" "a" "a" "c" "t" "a" "t" "a" "a" "a" "a"
  ...

> first_codon = Papio_hamadryas_vector[-c(4:length(Papio_hamadryas_vector))]

> first_codon
[1] "a" "t" "g"

> c2s(first_codon)
[1] "atg"


> P = Papio_hamadryas_vector


###############################################################################

> fcodons = function(x)
{
library(seqinr)
codons = NULL
for (i in 1:length(x))
{
codons[i] = c2s(x[-c(4:length(x))])
x = x[-c(1:3)]
}
codons = codons[1:(length(codons)/3)]
return(codons)
}

###############################################################################


> fcodons(P)
   [1] "atg" "caa" "aaa" "ctg" "caa" "ctc" "tgt" "gtt" "tat" "att" "tac" "ctg"
  [13] "ttt" "atg" "ctg" "att" "gtt" "gct" "ggt" "cca" "gtg" "gat" "cta" "aat"
  [25] "gag" "aac" "agt" "gag" "caa" "aaa" "gaa" "aat" "gtg" "gaa" "aaa" "gag"
  ...


> translate(P)
   [1] "M" "Q" "K" "L" "Q" "L" "C" "V" "Y" "I" "Y" "L" "F" "M" "L" "I" "V" "A"
  [19] "G" "P" "V" "D" "L" "N" "E" "N" "S" "E" "Q" "K" "E" "N" "V" "E" "K" "E"
  [37] "G" "L" "C" "N" "A" "C" "T" "W" "R" "Q" "N" "T" "K" "S" "S" "R" "I" "E"
  ...


> count_codons = count(P, word = 3)
> write.table(count_codons, file='tabella_codoni.xlsx')
```

```r
> codons_df = as.data.frame(count_codons)

> aa = c2s(translate(P))

> aa = AAString(aa)

> aa
376-letter AAString object
seq: MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKE...TPTKMSPINMLYFNGKEQIIYGKIPAMVVDRCGCS*

> AMINO_ACID_CODE[strsplit(as.character(aa), NULL)[[1]]]
    M     Q     K     L     Q     L     C     V     Y     I     Y     L     F
"Met" "Gln" "Lys" "Leu" "Gln" "Leu" "Cys" "Val" "Tyr" "Ile" "Tyr" "Leu" "Phe"
    M     L     I     V     A     G     P     V     D     L     N     E     N
"Met" "Leu" "Ile" "Val" "Ala" "Gly" "Pro" "Val" "Asp" "Leu" "Asn" "Glu" "Asn"
    S     E     Q     K     E     N     V     E     K     E     G     L     C
"Ser" "Glu" "Gln" "Lys" "Glu" "Asn" "Val" "Glu" "Lys" "Glu" "Gly" "Leu" "Cys"
    N     A     C     T     W     R     Q     N     T     K     S     S     R
"Asn" "Ala" "Cys" "Thr" "Trp" "Arg" "Gln" "Asn" "Thr" "Lys" "Ser" "Ser" "Arg"
    I     E     A     I     K     I     Q     I     L     S     K     L     R
"Ile" "Glu" "Ala" "Ile" "Lys" "Ile" "Gln" "Ile" "Leu" "Ser" "Lys" "Leu" "Arg"
    L     E     T     A     P     N     I     S     K     D     A     I     R
"Leu" "Glu" "Thr" "Ala" "Pro" "Asn" "Ile" "Ser" "Lys" "Asp" "Ala" "Ile" "Arg"
    Q     L     L     P     K     A     P     P     L     R     E     L     I
"Gln" "Leu" "Leu" "Pro" "Lys" "Ala" "Pro" "Pro" "Leu" "Arg" "Glu" "Leu" "Ile"
    D     Q     Y     D     V     Q     R     D     D     S     S     D     G
"Asp" "Gln" "Tyr" "Asp" "Val" "Gln" "Arg" "Asp" "Asp" "Ser" "Ser" "Asp" "Gly"
    S     L     E     D     D     D     Y     H     A     T     T     E     T
"Ser" "Leu" "Glu" "Asp" "Asp" "Asp" "Tyr" "His" "Ala" "Thr" "Thr" "Glu" "Thr"
    I     I     T     M     P     T     E     S     D     F     L     M     Q
"Ile" "Ile" "Thr" "Met" "Pro" "Thr" "Glu" "Ser" "Asp" "Phe" "Leu" "Met" "Gln"
    V     D     G     K     P     K     C     C     F     F     K     F     S
"Val" "Asp" "Gly" "Lys" "Pro" "Lys" "Cys" "Cys" "Phe" "Phe" "Lys" "Phe" "Ser"
    S     K     I     Q     Y     N     K     V     V     K     A     Q     L
"Ser" "Lys" "Ile" "Gln" "Tyr" "Asn" "Lys" "Val" "Val" "Lys" "Ala" "Gln" "Leu"
    W     I     Y     L     R     P     V     E     T     P     T     T     V
"Trp" "Ile" "Tyr" "Leu" "Arg" "Pro" "Val" "Glu" "Thr" "Pro" "Thr" "Thr" "Val"
    F     V     Q     I     L     R     L     I     K     P     M     K     D
"Phe" "Val" "Gln" "Ile" "Leu" "Arg" "Leu" "Ile" "Lys" "Pro" "Met" "Lys" "Asp"
    G     T     R     Y     T     G     I     R     S     L     K     L     D
"Gly" "Thr" "Arg" "Tyr" "Thr" "Gly" "Ile" "Arg" "Ser" "Leu" "Lys" "Leu" "Asp"
    M     N     P     G     T     G     I     W     Q     S     I     D     V
"Met" "Asn" "Pro" "Gly" "Thr" "Gly" "Ile" "Trp" "Gln" "Ser" "Ile" "Asp" "Val"
    K     T     V     L     Q     N     W     L     K     Q     P     E     S
"Lys" "Thr" "Val" "Leu" "Gln" "Asn" "Trp" "Leu" "Lys" "Gln" "Pro" "Glu" "Ser"
    N     L     G     I     E     I     K     A     L     D     E     N     G
"Asn" "Leu" "Gly" "Ile" "Glu" "Ile" "Lys" "Ala" "Leu" "Asp" "Glu" "Asn" "Gly"
    H     D     L     A     V     T     F     P     G     P     G     E     D
"His" "Asp" "Leu" "Ala" "Val" "Thr" "Phe" "Pro" "Gly" "Pro" "Gly" "Glu" "Asp"
    G     L     N     P     F     L     E     V     K     V     T     D     T
"Gly" "Leu" "Asn" "Pro" "Phe" "Leu" "Glu" "Val" "Lys" "Val" "Thr" "Asp" "Thr"
    P     K     R     S     R     R     D     F     G     L     D     C     D
"Pro" "Lys" "Arg" "Ser" "Arg" "Arg" "Asp" "Phe" "Gly" "Leu" "Asp" "Cys" "Asp"
    E     H     S     T     E     S     R     C     C     R     Y     P     L
"Glu" "His" "Ser" "Thr" "Glu" "Ser" "Arg" "Cys" "Cys" "Arg" "Tyr" "Pro" "Leu"
    T     V     D     F     E     A     L     G     W     D     W     I     I
"Thr" "Val" "Asp" "Phe" "Glu" "Ala" "Leu" "Gly" "Trp" "Asp" "Trp" "Ile" "Ile"
    A     P     K     R     Y     K     A     N     Y     C     S     G     E
"Ala" "Pro" "Lys" "Arg" "Tyr" "Lys" "Ala" "Asn" "Tyr" "Cys" "Ser" "Gly" "Glu"
    C     E     F     V     F     L     Q     K     Y     P     H     T     H
"Cys" "Glu" "Phe" "Val" "Phe" "Leu" "Gln" "Lys" "Tyr" "Pro" "His" "Thr" "His"
    L     V     H     Q     A     N     P     R     G     S     A     G     P
"Leu" "Val" "His" "Gln" "Ala" "Asn" "Pro" "Arg" "Gly" "Ser" "Ala" "Gly" "Pro"
    C     C     T     P     T     K     M     S     P     I     N     M     L
```

```
"Cys" "Cys" "Thr" "Pro" "Thr" "Lys" "Met" "Ser" "Pro" "Ile" "Asn" "Met" "Leu"
     Y     F     N     G     K     E     Q     I     I     Y     G     K     I
"Tyr" "Phe" "Asn" "Gly" "Lys" "Glu" "Gln" "Ile" "Ile" "Tyr" "Gly" "Lys" "Ile"
     P     A     M     V     V     D     R     C     G     C     S   <NA>
"Pro" "Ala" "Met" "Val" "Val" "Asp" "Arg" "Cys" "Gly" "Cys" "Ser"    NA
```

> **Papio_hamadryas_protein = c2s(translate(P))**
> **Papio_hamadryas_protein**
[1]
"MQKLQLCVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACTWRQNTKSSRIEAIKIQILSKLRLETAPNISKDAIRQLLPKAPPLR
ELIDQYDVQRDDSSDGSLEDDDYHATTETIITMPTESDFLMQVDGKPKCCFFKFSSKIQYNKVVKAQLWIYLRPVETPTTVFVQILRLI
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDENGHDLAVTFPGPGEDGLNPFLEVKVTDTPKRSRR
DFGLDCDEHSTESRCCRYPLTVDFEALGWDWIIAPKRYKANYCSGECEFVFLQKYPHTHLVHQANPRGSAGPCCTPTKMSPINMLYFNG
KEQIIYGKIPAMVVDRCGCS*"

> **bovine_belgian.fas = read.dna('bovine_belgian.fas', format = 'fasta')**
> **bovine_belgian.fas**
1 DNA sequence in binary format stored in a matrix.

Sequence length: 1128

Label:
Bovine belgian myostatin (MSTN) mRNA, complete cds

Base composition:
    a   c   g   t
0.312 0.212 0.222 0.254
(Total: 1.13 kb)

> **bovine_belgian_string = c2s(bovine_belgian.fas)**
> bovine_belgian_string
[1]
"atgcaaaaactgcaaatctctgtttatatttacctatttatgctgattgttgctggcccagtggatctgaatgagaacagcgagcaga
aggaaaatgtggaaaaagagggggctgtgtaatgcatgtttgtggagggaaaacactacatcctcaagactagaagccataaaaatccaa
atcctcagtaaacttcgcctggaaacagctcctaacatcagcaaagatgctatcagacaacttttgcccaaggctcctccactcctgga
actgattgatcagttcgatgtccagagagatgccagcagtgacggctccttggaagacgatgactaccacgccaggacggaaacggtca
ttaccatgcccacggagtctgatcttctaacgcaagtggaaggaaaacccaaatgttgcttctttaaatttagctctaagatacaatac
aataaactagtaaaggcccaactgtggatatatctgaggcctgtcaagactcctgcgacagtgtttgtgcaaatcctgagactcatcaa
acccatgaaagacggtacaaggtatactggaatccgatctctgaaacttgacatgaacccaggcactggtatttggcagagcattgatg
tgaagacagtgttgcagaactggctcaaacaacctgaatccaacttaggcattgaaatcaaagctttagatgagaatggccatgatctt
gctgtaaccttcccagaaccaggagaagatggactgactccttttttttagaagtcaaggtaacagacacaccaaaaagatctaggagaga
ttttgggcttgattgtga-----------
cagaatctcgatgctgtcgttaccctctaactgtggattttgaagcttttggatgggattggattattgcacctaaaagatataaggcc
aattactgctctggagaatgtgaatttgtattttttgcaaaagtatcctcatacccatcttgtgcaccaagcaaaccccagaggttcagc
cggcccctgctgtactcctacaaagatgtctccaattaatatgctatattttaatggcgaaggacaaataatatacgggaagattccag
ccatggtagtagatcgctgtgggtgttcatga"

> **translate(s2c(bovine_belgian_string))**
  [1] "M" "Q" "K" "L" "Q" "I" "S" "V" "Y" "I" "Y" "L" "F" "M" "L" "I" "V" "A"
 [19] "G" "P" "V" "D" "L" "N" "E" "N" "S" "E" "Q" "K" "E" "N" "V" "E" "K" "E"
 [37] "G" "L" "C" "N" "A" "C" "L" "W" "R" "E" "N" "T" "T" "S" "S" "R" "L" "E"
 [55] "A" "I" "K" "I" "Q" "I" "L" "S" "K" "L" "R" "L" "E" "T" "A" "P" "N" "I"
 [73] "S" "K" "D" "A" "I" "R" "Q" "L" "L" "P" "K" "A" "P" "P" "L" "L" "E" "L"
 [91] "I" "D" "Q" "F" "D" "V" "Q" "R" "D" "A" "S" "S" "D" "G" "S" "L" "E" "D"
[109] "D" "D" "Y" "H" "A" "R" "T" "E" "T" "V" "I" "T" "M" "P" "T" "E" "S" "D"
[127] "L" "L" "T" "Q" "V" "E" "G" "K" "P" "K" "C" "C" "F" "F" "K" "F" "S" "S"
[145] "K" "I" "Q" "Y" "N" "K" "L" "V" "K" "A" "Q" "L" "W" "I" "Y" "L" "R" "P"
[163] "V" "K" "T" "P" "A" "T" "V" "F" "V" "Q" "I" "L" "R" "L" "I" "K" "P" "M"
[181] "K" "D" "G" "T" "R" "Y" "T" "G" "I" "R" "S" "L" "K" "L" "D" "M" "N" "P"
[199] "G" "T" "G" "I" "W" "Q" "S" "I" "D" "V" "K" "T" "V" "L" "Q" "N" "W" "L"
[217] "K" "Q" "P" "E" "S" "N" "L" "G" "I" "E" "I" "K" "A" "L" "D" "E" "N" "G"
[235] "H" "D" "L" "A" "V" "T" "F" "P" "E" "P" "G" "E" "D" "G" "L" "T" "P" "F"
[253] "L" "E" "V" "K" "V" "T" "D" "T" "P" "K" "R" "S" "R" "R" "D" "F" "G" "L"
[271] "D" "C" "D" "R" "I" "S" "M" "L" "S" "L" "P" "S" "N" "C" "G" "F" "*" "S"

```
[289] "F" "W" "M" "G" "L" "D" "Y" "C" "T" "*" "K" "I" "*" "G" "Q" "L" "L" "L"
[307] "W" "R" "M" "*" "I" "C" "I" "F" "A" "K" "V" "S" "S" "Y" "P" "S" "C" "A"
[325] "P" "S" "K" "P" "Q" "R" "F" "S" "R" "P" "L" "L" "Y" "S" "Y" "K" "D" "V"
[343] "S" "N" "*" "Y" "A" "I" "F" "*" "W" "R" "R" "T" "N" "N" "I" "R" "E" "D"
[361] "S" "S" "H" "G" "S" "R" "S" "L" "W" "V" "F" "M"
```

```
> bovine_belgian_protein = c2s(translate(s2c(bovine_belgian_string)))
> bovine_belgian_protein
[1]
"MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSKLRLETAPNISKDAIRQLLPKAPPLL
ELIDQFDVQRDASSDGSLEDDDYHARTETVITMPTESDLLTQVEGKPKCCFFKFSSKIQYNKLVKAQLWIYLRPVKTPATVFVQILRLI
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDENGHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRR
DFGLDCXXXXXESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFLQKYPHTHLVHQANPRGSAGPCCTPTKMSPINMLYFNG
EGQIIYGKIPAMVVDRCGCS*"
```

```
> c2s(translate(s2c(bovine_belgian_string)))
[1]
"MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSKLRLETAPNISKDAIRQLLPKAPPLL
ELIDQFDVQRDASSDGSLEDDDYHARTETVITMPTESDLLTQVEGKPKCCFFKFSSKIQYNKLVKAQLWIYLRPVKTPATVFVQILRLI
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDENGHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRR
DFGLDCXXXXXESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFLQKYPHTHLVHQANPRGSAGPCCTPTKMSPINMLYFNG
EGQIIYGKIPAMVVDRCGCS*"
> aa = c2s(translate(s2c(bovine_belgian_string)))
> aa
[1]
"MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSKLRLETAPNISKDAIRQLLPKAPPLL
ELIDQFDVQRDASSDGSLEDDDYHARTETVITMPTESDLLTQVEGKPKCCFFKFSSKIQYNKLVKAQLWIYLRPVKTPATVFVQILRLI
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDENGHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRR
DFGLDCXXXXXESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGECEFVFLQKYPHTHLVHQANPRGSAGPCCTPTKMSPINMLYFNG
EGQIIYGKIPAMVVDRCGCS*"
> AMINO_ACID_CODE[strsplit(as.character(aa), NULL)[[1]]]
    M     Q     K     L     Q     I     S     V     Y     I     Y     L     F
"Met" "Gln" "Lys" "Leu" "Gln" "Ile" "Ser" "Val" "Tyr" "Ile" "Tyr" "Leu" "Phe"
    M     L     I     V     A     G     P     V     D     L     N     E     N
"Met" "Leu" "Ile" "Val" "Ala" "Gly" "Pro" "Val" "Asp" "Leu" "Asn" "Glu" "Asn"
    S     E     Q     K     E     N     V     E     K     E     G     L     C
"Ser" "Glu" "Gln" "Lys" "Glu" "Asn" "Val" "Glu" "Lys" "Glu" "Gly" "Leu" "Cys"
    N     A     C     L     W     R     E     N     T     T     S     S     R
"Asn" "Ala" "Cys" "Leu" "Trp" "Arg" "Glu" "Asn" "Thr" "Thr" "Ser" "Ser" "Arg"
    L     E     A     I     K     I     Q     I     L     S     K     L     R
"Leu" "Glu" "Ala" "Ile" "Lys" "Ile" "Gln" "Ile" "Leu" "Ser" "Lys" "Leu" "Arg"
    L     E     T     A     P     N     I     S     K     D     A     I     R
"Leu" "Glu" "Thr" "Ala" "Pro" "Asn" "Ile" "Ser" "Lys" "Asp" "Ala" "Ile" "Arg"
    Q     L     L     P     K     A     P     P     L     L     E     L     I
"Gln" "Leu" "Leu" "Pro" "Lys" "Ala" "Pro" "Pro" "Leu" "Leu" "Glu" "Leu" "Ile"
    D     Q     F     D     V     Q     R     D     A     S     S     D     G
"Asp" "Gln" "Phe" "Asp" "Val" "Gln" "Arg" "Asp" "Ala" "Ser" "Ser" "Asp" "Gly"
    S     L     E     D     D     D     Y     H     A     R     T     E     T
"Ser" "Leu" "Glu" "Asp" "Asp" "Asp" "Tyr" "His" "Ala" "Arg" "Thr" "Glu" "Thr"
    V     I     T     M     P     T     E     S     D     L     L     T     Q
"Val" "Ile" "Thr" "Met" "Pro" "Thr" "Glu" "Ser" "Asp" "Leu" "Leu" "Thr" "Gln"
    V     E     G     K     P     K     C     C     F     F     K     F     S
"Val" "Glu" "Gly" "Lys" "Pro" "Lys" "Cys" "Cys" "Phe" "Phe" "Lys" "Phe" "Ser"
    S     K     I     Q     Y     N     K     L     V     K     A     Q     L
"Ser" "Lys" "Ile" "Gln" "Tyr" "Asn" "Lys" "Leu" "Val" "Lys" "Ala" "Gln" "Leu"
    W     I     Y     L     R     P     V     K     T     P     A     T     V
"Trp" "Ile" "Tyr" "Leu" "Arg" "Pro" "Val" "Lys" "Thr" "Pro" "Ala" "Thr" "Val"
    F     V     Q     I     L     R     L     I     K     P     M     K     D
"Phe" "Val" "Gln" "Ile" "Leu" "Arg" "Leu" "Ile" "Lys" "Pro" "Met" "Lys" "Asp"
    G     T     R     Y     T     G     I     R     S     L     K     L     D
"Gly" "Thr" "Arg" "Tyr" "Thr" "Gly" "Ile" "Arg" "Ser" "Leu" "Lys" "Leu" "Asp"
    M     N     P     G     T     G     I     W     Q     S     I     D     V
"Met" "Asn" "Pro" "Gly" "Thr" "Gly" "Ile" "Trp" "Gln" "Ser" "Ile" "Asp" "Val"
    K     T     V     L     Q     N     W     L     K     Q     P     E     S
```

```
"Lys" "Thr" "Val" "Leu" "Gln" "Asn" "Trp" "Leu" "Lys" "Gln" "Pro" "Glu" "Ser"
   N     L     G     I     E     I     K     A     L     D     E     N     G
"Asn" "Leu" "Gly" "Ile" "Glu" "Ile" "Lys" "Ala" "Leu" "Asp" "Glu" "Asn" "Gly"
   H     D     L     A     V     T     F     P     E     P     G     E     D
"His" "Asp" "Leu" "Ala" "Val" "Thr" "Phe" "Pro" "Glu" "Pro" "Gly" "Glu" "Asp"
   G     L     T     P     F     L     E     V     K     V     T     D     T
"Gly" "Leu" "Thr" "Pro" "Phe" "Leu" "Glu" "Val" "Lys" "Val" "Thr" "Asp" "Thr"
   P     K     R     S     R     R     D     F     G     L     D     C     X
"Pro" "Lys" "Arg" "Ser" "Arg" "Arg" "Asp" "Phe" "Gly" "Leu" "Asp" "Cys" "Xaa"
   X     X     X     X     E     S     R     C     C     R     Y     P     L
"Xaa" "Xaa" "Xaa" "Xaa" "Glu" "Ser" "Arg" "Cys" "Cys" "Arg" "Tyr" "Pro" "Leu"
   T     V     D     F     E     A     F     G     W     D     W     I     I
"Thr" "Val" "Asp" "Phe" "Glu" "Ala" "Phe" "Gly" "Trp" "Asp" "Trp" "Ile" "Ile"
   A     P     K     R     Y     K     A     N     Y     C     S     G     E
"Ala" "Pro" "Lys" "Arg" "Tyr" "Lys" "Ala" "Asn" "Tyr" "Cys" "Ser" "Gly" "Glu"
   C     E     F     V     F     L     Q     K     Y     P     H     T     H
"Cys" "Glu" "Phe" "Val" "Phe" "Leu" "Gln" "Lys" "Tyr" "Pro" "His" "Thr" "His"
   L     V     H     Q     A     N     P     R     G     S     A     G     P
"Leu" "Val" "His" "Gln" "Ala" "Asn" "Pro" "Arg" "Gly" "Ser" "Ala" "Gly" "Pro"
   C     C     T     P     T     K     M     S     P     I     N     M     L
"Cys" "Cys" "Thr" "Pro" "Thr" "Lys" "Met" "Ser" "Pro" "Ile" "Asn" "Met" "Leu"
   Y     F     N     G     E     G     Q     I     I     Y     G     K     I
"Tyr" "Phe" "Asn" "Gly" "Glu" "Gly" "Gln" "Ile" "Ile" "Tyr" "Gly" "Lys" "Ile"
   P     A     M     V     V     D     R     C     G     C     S    <NA>
"Pro" "Ala" "Met" "Val" "Val" "Asp" "Arg" "Cys" "Gly" "Cys" "Ser"    NA
```

> **bovine_piedmontese.fas = read.dna('bovine_piedmontese.fas', format = 'fasta')**

> **bovine_piedmontese.fas**
1 DNA sequence in binary format stored in a matrix.

Sequence length: 1128

Label:
Bovie_piedmotese

Base composition:
    a   c   g   t
0.313 0.214 0.220 0.254
(Total: 1.13 kb)

> **bovine_piedmontese_string = c2s(bovine_piedmontese.fas)**

> **bovine_piedmontese_string**
[1]
"atgcaaaaactgcaaatctctgtttatatttacctatttatgctgattgttgctggcccagtggatctgaatgagaacagcgagcaga
aggaaaatgtggaaaaagaggggctgtgtaatgcatgtttgtggagggaaaacactacatcctcaagactagaagccataaaaatccaa
atcctcagtaaacttcgcctggaaacagctcctaacatcagcaaagatgctatcagacaacttttgcccaaggctcctccactcctgga
actgattgatcagttcgatgtccagagagatgccagcagtgacggctccttggaagacgatgactaccacgccaggacggaaacggtca
ttaccatgcccacggagtctgatcttctaacgcaagtggaaggaaaacccaaatgttgcttctttaaatttagctctaagatacaatac
aataaactagtaaaggcccaactgtggatatatctgaggcctgtcaagactcctgcgacagtgtttgtgcaaatcctgagactcatcaa
acccatgaaagacggtacaaggtatactggaatccgatctctgaaacttgacatgaacccaggcactggtatttggcagagcattgatg
tgaagacagtgttgcagaactggctcaaacaacctgaatccaacttaggcattgaaatcaaagctttagatgagaatggccatgatctt
gctgtaaccttcccagaaccaggagaagatggactgactccttttttagaagtcaaggtaacagacacaccaaaaagatctaggagaga
ttttgggcttgattgtgatgaacactccacagaatctcgatgctgtcgttaccctctaactgtggattttgaagctttggatggggatt
ggattattgcacctaaaagatataaggccaattactgctctggagaatatgaatttgtattttgcaaaagtatcctcatacccatcttt
gtgcaccaagcaaaccccagaggttcagccggcccctgctgtactcctacaaagatgtctccaattaatatgctatatttttaatggcga
aggacaaataatatacgggaagattccagccatggtagtagatcgctgtgggtgttcatga"

> **aa = c2s(translate(s2c(bovine_piedmontese_string)))**

> aa
[1]
"MQKLQISVYIYLFMLIVAGPVDLNENSEQKENVEKEGLCNACLWRENTTSSRLEAIKIQILSKLRLETAPNISKDAIRQLLPKAPPLL
```

ELIDQFDVQRDASSDGSLEDDDYHARTETVITMPTESDLLTQVEGKPKCCFFKFSSKIQYNKLVKAQLWIYLRPVKTPATVFVQILRLI
KPMKDGTRYTGIRSLKLDMNPGTGIWQSIDVKTVLQNWLKQPESNLGIEIKALDENGHDLAVTFPEPGEDGLTPFLEVKVTDTPKRSRR
DFGLDCDEHSTESRCCRYPLTVDFEAFGWDWIIAPKRYKANYCSGEYEFVFLQKYPHTHLVHQANPRGSAGPCCTPTKMSPINMLYFNG
EGQIIYGKIPAMVVDRCGCS*"

```
> Bos_taurus.fas = read.dna('Bos_taurus.fas', format = 'fasta')

> Bos_taurus_string = c2s(Bos_taurus.fas)

> Bos_taurus_string
[1]
"atgcaaaaactgcaaatctctgtttatatttacctatttatgctgattgttgctggcccagtggatctgaatgagaacagcgagcaga
aggaaaatgtggaaaaagaggggctgtgtaatgcatgtttgtggagggaaaacactacatcctcaagactagaagccataaaaatccaa
atcctcagtaaacttcgcctggaaacagctcctaacatcagcaaagatgctatcagacaacttttgcccaaggctcctccactcctgga
actgattgatcagttcgatgtccagagagatgccagcagtgacggctccttggaagacgatgactaccacgccaggacggaaacggtca
ttaccatgcccacggagtctgatcttctaacgcaagtggaaggaaaacccaaatgttgcttctttaaatttagctctaagatacaatac
aataaactagtaaaggcccaactgtggatatatctgaggcctgtcaagactcctgcgacagtgtttgtgcaaatcctgagactcatcaa
acccatgaaagacggtacaaggtatactggaatccgatctctgaaacttgacatgaacccaggcactggtatttggcagagcattgatg
tgaagacagtgttgcagaactggctcaaacaacctgaatccaacttaggcattgaaatcaaagctttagatgagaatggccatgatctt
gctgtaaccttcccagaaccaggagaagatggactgactcctttttttagaagtcaaggtaacagacacaccaaaaagatctaggagaga
ttttgggcttgattgtgatgaacactccacagaatctcgatgctgtcgttaccctctaactgtggattttgaagcttttggatgggatt
ggattattgcacctaaaagatataaggccaattactgctctggagaatgtgaatttgtattttttgcaaaagtatcctcatacccatctt
gtgcaccaagcaaaccccagaggttcagccggcccctgctgtactcctacaaagatgtctccaattaatatgctatattttaatggcga
aggacaaataatatacgggaagattccagccatggtagtagatcgctgtgggtgttcatga"

> aa = c2s(translate(s2c(Bos_taurus_string)))

> save.image(file='codons.Rdata')
```

```
> codons_df
   Var1 Freq
1   aaa   55
2   aac   26
3   aag   25
...
```