

Review - DNA or RNA

Luna L Sanchez Reyes

2022-12-08

```
read.csv(file = "uhuru") <- min(1:10)
dna_or_rna(sequences)
my_name <- "Luna"
anything_any_name_you_want <- function()
```

1. Write a function, `dna_or_rna(sequence)`, that determines if a sequence of base pairs is DNA, RNA, or if it is not possible to tell given the sequence provided. Since all the function will know about the material is the sequence, the only way to tell the difference between DNA and RNA is that RNA has the base Uracil (“u”) instead of the base Thymine (“t”). Have the function return one of three outputs: “DNA”, “RNA”, or “UNKNOWN”.

```
"a" == "a"
```

```
## [1] TRUE
```

```
"a" == "b"
```

```
## [1] FALSE
```

```
"t" == "actgc"
```

```
## [1] FALSE
```

```
is.element("a", "a")
```

```
## [1] TRUE
```

```
is.element("a", "actgc")
```

```
## [1] FALSE
```

```
is.element("a", c("acgtc", "a", "tttgc"))
```

```
## [1] TRUE
```

```
str_detect("a", "actgc")
```

```
## [1] FALSE
```

```
str_detect("agtc", "a")
```

```
## [1] TRUE
```

```
?str_detect
```

```
library(stringr)
```

```
dna_or_rna <- function(sequence){
  if (str_detect(string = sequence, "u")) {
    print("RNA")
  } else if (str_detect(string = sequence, pattern = "t")) {
```

```

    print("DNA")
  } else {
    print("unknown")
  }
}

```

```

sequences = c("ttgaatgccttacaactgatcattacacaggcggcatgaagcaaaaatatactgtgaaccaatgcaggcg", "gauuauuccccacaaaggaggagugggauuaggagcugcaucauuuacaagagcagaauuucaaaugcau", "gaaagcaagaaaaggcaggcgaggaagggaagaaggggggaaacc", "guuuccuacaguauuugaugagagaugagaguuuacuccuggaagauaaauuagaauuuuacaacugcaccugaucagguggauaaggaagaugac", "gauaaggaagaugaagacuucaggaaucauaauaaaugcacuccaugaauaggauucauguaugggaaucagccggguc")

```

2. Use the function and a for loop to print the type of the sequences in the list above.

```

for (i in sequences) {
  dna_or_rna(i)
  print(i)
}

```

```

## [1] "DNA"
## [1] "ttgaatgccttacaactgatcattacacaggcggcatgaagcaaaaatatactgtgaaccaatgcaggcg"
## [1] "RNA"
## [1] "gauuauuccccacaaaggaggagugggauuaggagcugcaucauuuacaagagcagaauuucaaaugcau"
## [1] "unknown"
## [1] "gaaagcaagaaaaggcaggcgaggaagggaagaaggggggaaacc"
## [1] "RNA"
## [1] "guuuccuacaguauuugaugagagaugagaguuuacuccuggaagauaaauuagaauuuuacaacugcaccugaucagguggauaaggaagaugac"
## [1] "RNA"
## [1] "gauaaggaagaugaagacuucaggaaucauaauaaaugcacuccaugaauaggauucauguaugggaaucagccggguc"

```

3. Use the function and sapply to print the type of the sequences in the list below.

```

sapply(X = sequences, FUN = dna_or_rna)

```

```

## [1] "DNA"
## [1] "RNA"
## [1] "unknown"
## [1] "RNA"
## [1] "RNA"

##                               ttgaatgccttacaactgatcattacacaggcggcatgaagcaaaaatatactgtgaaccaatgcaggcg
##                               "DNA"
##                               gauuauuccccacaaaggaggagugggauuaggagcugcaucauuuacaagagcagaauuucaaaugcau
##                               "RNA"
##                               gaaagcaagaaaaggcaggcgaggaagggaagaaggggggaaacc
##                               "unknown"
## guuuccuacaguauuugaugagagaugagaguuuacuccuggaagauaaauuagaauuuuacaacugcaccugaucagguggauaaggaagaugac
##                               "RNA"
##                               gauaaggaagaugaagacuucaggaaucauaauaaaugcacuccaugaauaggauucauguaugggaaucagccggguc
##                               "RNA"

```

4. Optional: For a little extra challenge make your function work with both upper and lower case letters, or even strings with mixed capitalization

sequ

```

sequences <- c("ATGC",
              "UCCCCGU")
str_to_lower(sequences)

```

```

## [1] "atgc"      "uccccgu"

```

```

tolower(sequences)

## [1] "atgc"      "uccccgu"
dna_or_rna(sequences[1])

## [1] "unknown"
dna_or_rna <- function(sequence){
  sequence <- tolower(sequence)
  #sequence <- str_to_lower(sequence)
  if (str_detect(string = sequence, c("u"))) {
    print("RNA")
  } else if (str_detect(string = sequence, pattern = "t")) {
    print("DNA")
  } else {
    print("unknown")
  }
}

#dna_or_rna <- function(sequence){
#  sequence <- tolower(sequence)
#  if (str_detect(string = sequence, c("u")) | str_detect(string = sequence, #pattern = "U")) {
#    print("RNA")
#  } else if (str_detect(string = sequence, pattern = "U")) {
#    print("RNA")
#  } else if (str_detect(string = sequence, pattern = "t")) {
#    print("DNA")
#  } else {
#    print("unknown")
#  }
#}
dna_or_rna(sequences[1])

## [1] "DNA"

```

5. Elaborate on the main difference between a for loop and an sapply function: