

bio_R

2025-05-12

Introduction

Package contains 3 functions:

- reverse complement computation,
- one-hot encoding and
- GC content plotting.

Package download

```
if (!require(ggplot2)) install.packages("ggplot2")
library(ggplot2)
```

Data loading

```
load("example_sequences.RData")
head(example_sequences)
```

Example data

```
data.frame(ID = names(example_sequences)[1:5],
           Sequence = unname(example_sequences[1:5]))
```

seq_id	sequence
seq_1	GGGCGCCCGATCCACGTAGG
seq_2	ATAAAGTCGCACGTCAGGAT
seq_3	GTCAGAACGGTAGAGTCACA
seq_4	ATTGACAAGACAGAGCTGTT
seq_5	CCGTCCGGTACCACAAAAAA

Reverse complement and one-hot encoding

```
seq1 <- example_sequences[[1]]
rc <- reverse_complement_dna(seq1)
encoded <- one_hot_encode_dna(rc)
encoded
```

GC content plot

```
plot_gc_content(example_sequences)
```

