SECB 3213-01
BIOINFORMATICS DATABASE
SEMESTER 2 2022/2023

ASSIGNMENT 1

LECTURER:
DR NOOR HIDAYAH ZAKARIA

SUBMITTED BY:
CHONG KAH WEI (A20EC0027)

Screenshot of hands-on

```
> #Main Pairwise Sequence Alignment Function
> library(Biostrings)
> pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede")
Global PairwiseAlignmentsSingleSubject (1 of 2)
pattern: succ--eed
subject: supersede
score: -33.99738
> pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede", type = "local")
Local PairwiseAlignmentsSingleSubject (1 of 2)
pattern: [1] su
subject: [1] su
score: 5.578203
> pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede", gapOpening = 0, gapExtension = 1)
Global PairwiseAlignmentsSingleSubject (1 of 2)
pattern: su-cce--ed-
subject: sup--ersede
score: 7.945507
> submat <- matrix(-1, nrow = 26, ncol = 26, dimnames = list(letters, letters))
> diag(submat) <- 0
> pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede",substitutionMatrix = submat,gapopening = 0, gapExtension = 1)
Global PairwiseAlignmentsSingleSubject (1 of 2)
pattern: succe-ed-
subject: supersede
score: -5
> submat <- matrix(-1, nrow = 26, ncol = 26, dimnames = list(letters,letters))
> diag(submat) <- 0
> pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede", substitutionMatrix = submat,gapopening = 0, gapExtension = 1, scoreonly = TRUE)
[1] -5 -5
```

Exercise 1

```
> #Exercise 1
> #1
> #global
> pairwiseAlignment("syzygy", "zyzzyx", type = "global")
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: syzygy
subject: zyzzyx
score: -19.3607
> #local
> pairwiseAlignment("syzygy", "zyzzyx", type = "local")
Local PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [3] zy
subject: [1] zy
score: 4.607359
> #overlap
> pairwiseAlignment("syzygy", "zyzzyx", type = "overlap")
Overlap PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [1]
subject: [7]
score: 0
> #2
> pairwiseAlignment("syzygy", "zyzzyx", type = "global", gapExtension = -Inf)
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: syzyg-y
subject: zyzzyx-
score: -Inf
> #local
> pairwiseAlignment("syzygy", "zyzzyx", type = "local", gapExtension = -Inf)
Local PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [3] zy
subject: [1] zy
score: 4.607359
> #overlap
> pairwiseAlignment("syzygy", "zyzzyx", type = "overlap", gapExtension = -Inf)
Overlap PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [1]
subject: [7]
score: 0
```

1.
```
> #global
> pairwiseAlignment("syzygy", "zyzzyx", type = "global")
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: syzygy
subject: zyzzyx
score: -19.3607
> #local
> pairwiseAlignment("syzygy", "zyzzyx", type = "local")
Local PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [3] zy
subject: [1] zy
score: 4.607359
> #overlap
> pairwiseAlignment("syzygy", "zyzzyx", type = "overlap")
Overlap PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [1]
subject: [7]
score: 0
```

2.
```
> pairwiseAlignment("syzygy", "zyzzyx", type = "global", gapExtension =
 -Inf)
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: syzyg-y
subject: zyzzyx-
score: -Inf
> #local
> pairwiseAlignment("syzygy", "zyzzyx", type = "local", gapExtension =
-Inf)
Local PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [3] zy
subject: [1] zy
score: 4.607359
> #overlap
> pairwiseAlignment("syzygy", "zyzzyx", type = "overlap", gapExtension
= -Inf)
Overlap PairwiseAlignmentsSingleSubject (1 of 1)
pattern: [1]
subject: [7]
score: 0
```

Ans:

Yes, the global pairwise alignment changes whereby,

Result for gapExtension = default:
```
pattern: syzygy
subject: zyzzyx
score: -19.3607
```

Result for gapExtension = -Inf:
```
pattern: syzyg-y
subject: zyzzyx-
score: -Inf
```

Screenshot of hands-on

```
> #Pairwise sequence Alignment classes
> pa1 <- pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede")
> class(pa1)
[1] "PairwiseAlignmentsSingleSubject"
attr(,"package")
[1] "Biostrings"
> summary(pa1)
Global Single Subject Pairwise Alignments
Number of Alignments:  2

scores:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 -34.00  -31.78  -29.56  -29.56  -27.34  -25.12

Number of matches:
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   3.00    3.25    3.50    3.50    3.75    4.00

Top 7 mismatch counts:
  SubjectPosition Subject Pattern Count Probability
1               3       p       c     1         0.5
2               4       e       c     1         0.5
3               4       e       r     1         0.5
4               5       r       e     1         0.5
5               6       s       c     1         0.5
6               8       d       e     1         0.5
7               9       e       d     1         0.5
> class(summary(pa1))
[1] "PairwiseAlignmentsSingleSubjectSummary"
attr(,"package")
[1] "Biostrings"
> class(pattern(pa1))
[1] "QualityAlignedXStringSet"
attr(,"package")
[1] "Biostrings"
> submat <- matrix(-1, nrow = 26, ncol = 26, dimnames = list(letters, letters))
> diag(submat) <- 0
> pa2 <- pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede",substitutionMatrix = submat, gapOpening = 0, gapExtension = 1)
> class(pattern(pa2))
[1] "AlignedXStringSet"
attr(,"package")
[1] "Biostrings"
> #Pairwise sequence Alignment Helper Functions
> submat <- matrix(-1, nrow = 26, ncol = 26, dimnames = list(letters, letters))
> diag(submat) <- 0
> pa2 <-pairwiseAlignment(pattern = c("succeed", "precede"), subject = "supersede",substitutionMatrix = submat, gapOpening = 0, gapExtension = 1)
> score(pa2)
[1] -5 -5
> nedit(pa2)
[1] 4 5
> nmatch(pa2)
[1] 4 4
> nmismatch(pa2)
[1] 3 3
> nchar(pa2)
[1] 8 9
```

```
> aligned(pa2)
BStringSet object of length 2:
    width seq
[1]     9 succe-ed-
[2]     9 pr-ec-ede
> as.character(pa2)
[1] "succe-ed-" "pr-ec-ede"
> as.matrix(pa2)
     [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]
[1,] "s"  "u"  "c"  "c"  "e"  "-"  "e"  "d"  "-"
[2,] "p"  "r"  "-"  "e"  "c"  "-"  "e"  "d"  "e"
> #Application: Using Evolutionary Models in Protein Alignments
> data(BLOSUM50)
> BLOSUM50[1:4,1:4]
   A  R  N  D
A  5 -2 -1 -2
R -2  7 -1 -2
N -1 -1  7  2
D -2 -2  2  8
> nwdemo <-pairwiseAlignment(AAString("PAWHEAE"), AAString("HEAGAWGHEE"), substitutionMatrix = BLOSUM50, gapOpening = 0, gapExtension = 8)
> nwdemo
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: -PA--W-HEAE
subject: HEAGAWGHE-E
score: 1
> compareStrings(nwdemo)
[1] "?A--W-HE+E"
> pid(nwdemo)
[1] 50
```

Exercise 2

1.

```
> #Exercise2
> #1
> data(BLOSUM62)
> pairwiseAlignment(AAString("PAWHEAE"), AAString("HEAGAWGHEE"), substitutionMatrix = BLOSUM62, gapOpening = 12, gapExtension = 4)
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: P---AWHEAE
subject: HEAGAWGHEE
score: -9
```

```
> data(BLOSUM62)
> pairwiseAlignment(AAString("PAWHEAE"), AAString("HEAGAWGHEE"), substi
tutionMatrix = BLOSUM62, gapOpening = 12, gapExtension = 4)
Global PairwiseAlignmentsSingleSubject (1 of 1)
pattern: P---AWHEAE
subject: HEAGAWGHEE
score: -9
```

2. Increase of the value of the gap opening penalty makes the gaps less frequent and decrease of the value of gap extension penalty make the gaps shorter. These cause the alignment to change.