Házi feladat

Bevezetés a bioinformatikába - NBVBB1HBNE/BB1_EA_VÁL

Burian Sándor | AWXYHE

Kézzel:

BLOSUM

1	B	A	B	A
2	A	A	A	C
3	A	A	C	C
4	A	A	B	A
5	A	A	C	C
6	A	A	B	C

Gyakoriság

A	14
B	4
C	6
Σ	24

$q_{48} = (14/24 \cdot 14/24)/2$

Páronkénti gyakoriság

A A	26	→ 26/60 =
A B	8	→ 8/60 =
A C	10	→ 10/60 =
B B	3	→ 3/60 =
B C	6	→ 6/60 =
C C	7	→ 7/60 =

Σ   60

Valódi páronkénti gyakoriság

0.4(3)
0.1(3)
0.1(6)
0.05
0.1
0.11(6)

Várt páronkénti gyakoriság

A A	0.68
A B	0.19(6)
A C	0.291(6)
B B	0.0(5)
B C	0.8(3)
C C	0.125

$2 \log_2 \left(\dfrac{\text{valódi}}{\text{várt}} \right)$

A A	−1.302
A B	−1.0892
A C	−1.6152
B B	−3.203
B C	−6.1166
C C	−0.2000...

Kódolva:

```python
1.  from Bio import Align
2.  from Bio.Align import substitution_matrices
3.  from Bio.Align import MultipleSeqAlignment
4.  from Bio.SeqRecord import SeqRecord
5.  from Bio.Seq import Seq
6.
7.  aligner = Align.PairwiseAligner()
8.  aligner.substitution_matrix = substitution_matrices.load("BLOSUM62")
9.  sequencies = ["BABA", "AAAC", "AACC", "AABA", "AACC", "AABC"]
10. for i in sequencies:
11.     for j in sequencies:
12.         alignments = aligner.align(i, j)
13.         alignments = list(alignments)
14.         print("Number of alignments: %d" % len(alignments))
15.         alignment = alignments[0]
16.         print("Score = %.1f" % alignment.score)
17.         print(alignment)
18.
```

Kimenet:

```
1.  Biopython v1.79
2.  Number of alignments: 1
3.  Score = 16.0
4.  BABA
5.  ||||
6.  BABA
7.
8.  Number of alignments: 5
9.  Score = 8.0
10. BABA--
11. -|-|--
12. -A-AAC
13.
14. Number of alignments: 1
15. Score = 8.0
16. BABA--
17. -|-|--
18. -A-ACC
19.
20. Number of alignments: 3
21. Score = 12.0
22. BA-BA
23. -|-||
24. -AABA
25.
26. Number of alignments: 1
27. Score = 8.0
28. BABA--
29. -|-|--
30. -A-ACC
31.
32. Number of alignments: 10
33. Score = 8.0
34. BABA--
35. -|-|--
36. -A-ABC
```

```
 37.
 38. Number of alignments: 5
 39. Score = 8.0
 40. -A-AAC
 41. -|-|--
 42. BABA--
 43.
 44. Number of alignments: 1
 45. Score = 21.0
 46. AAAC
 47. ||||
 48. AAAC
 49.
 50. Number of alignments: 8
 51. Score = 17.0
 52. AAAC-
 53. ||-|-
 54. AA-CC
 55.
 56. Number of alignments: 1
 57. Score = 12.0
 58. AA-AC
 59. ||-|-
 60. AABA-
 61.
 62. Number of alignments: 8
 63. Score = 17.0
 64. AAAC-
 65. ||-|-
 66. AA-CC
 67.
 68. Number of alignments: 4
 69. Score = 17.0
 70. AAA-C
 71. ||--|
 72. AA-BC
 73.
 74. Number of alignments: 1
 75. Score = 8.0
 76. -A-ACC
 77. -|-|--
 78. BABA--
 79.
 80. Number of alignments: 8
 81. Score = 17.0
 82. AA-CC
 83. ||-|-
 84. AAAC-
 85.
 86. Number of alignments: 1
 87. Score = 26.0
 88. AACC
 89. ||||
 90. AACC
 91.
 92. Number of alignments: 11
 93. Score = 8.0
 94. AACC--
 95. ||----
 96. AA--BA
 97.
 98. Number of alignments: 1
 99. Score = 26.0
100.    AACC
```

```
101.    ||||
102.    AACC
103.
104.    Number of alignments: 3
105.    Score = 17.0
106.    AA-CC
107.    ||-|-
108.    AABC-
109.
110.    Number of alignments: 3
111.    Score = 12.0
112.    -AABA
113.    -|-||
114.    BA-BA
115.
116.    Number of alignments: 1
117.    Score = 12.0
118.    AABA-
119.    ||-|-
120.    AA-AC
121.
122.    Number of alignments: 11
123.    Score = 8.0
124.    AABA--
125.    ||----
126.    AA--CC
127.
128.    Number of alignments: 1
129.    Score = 16.0
130.    AABA
131.    ||||
132.    AABA
133.
134.    Number of alignments: 11
135.    Score = 8.0
136.    AABA--
137.    ||----
138.    AA--CC
139.
140.    Number of alignments: 3
141.    Score = 12.0
142.    AABA-
143.    |||--
144.    AAB-C
145.
146.    Number of alignments: 1
147.    Score = 8.0
148.    -A-ACC
149.    -|-|--
150.    BABA--
151.
152.    Number of alignments: 8
153.    Score = 17.0
154.    AA-CC
155.    ||-|-
156.    AAAC-
157.
158.    Number of alignments: 1
159.    Score = 26.0
160.    AACC
161.    ||||
162.    AACC
163.
164.    Number of alignments: 11
```

```
165.    Score = 8.0
166.    AACC--
167.    ||----
168.    AA--BA
169.
170.    Number of alignments: 1
171.    Score = 26.0
172.    AACC
173.    ||||
174.    AACC
175.
176.    Number of alignments: 3
177.    Score = 17.0
178.    AA-CC
179.    ||-|-
180.    AABC-
181.
182.    Number of alignments: 10
183.    Score = 8.0
184.    -AABC-
185.    -|-|--
186.    BA-B-A
187.
188.    Number of alignments: 4
189.    Score = 17.0
190.    AAB-C
191.    ||--|
192.    AA-AC
193.
194.    Number of alignments: 3
195.    Score = 17.0
196.    AABC-
197.    ||-|-
198.    AA-CC
199.
200.    Number of alignments: 3
201.    Score = 12.0
202.    AABC-
203.    |||--
204.    AAB-A
205.
206.    Number of alignments: 3
207.    Score = 17.0
208.    AABC-
209.    ||-|-
210.    AA-CC
211.
212.    Number of alignments: 1
213.    Score = 21.0
214.    AABC
215.    ||||
216.    AABC
217.
218.
```

Elérhető: https://www.kaggle.com/sndorburian/bioinfo-homework-01-v001