

```
var c = "GAAGTGAGT, TTCGTAAGT, AAGGTACTT, CTGGTGAGC, " +  
"AGAGTGAGT, CAGGTAGAG, ACTGTACGT, CTGGTGAGT, " +  
"TATGTAAGT, CGGGTGAGC";
```

```
var t = "AAAAAACAG, AAAAACAGG, AAAACAGGT, AAACAGGTG, " +  
"AACAGGTGA, ACAGGTGAG, CAGGTGAGT, AGGTGAGTA, " +  
"GGTGAGTAA, GTGAGTAAA, TGAGTAAAA, GAGTAAAAA, " +  
"AGTAAAAAA, GTAAAAAA";
```

(a)

rows(sp) = 10

The motif model

	1	2	3	4	5	6	7	8	9
1	G	A	A	G	T	G	A	G	T
2	T	T	C	G	T	A	A	G	T
3	A	A	G	G	T	A	C	T	T
4	C	T	G	G	T	G	A	G	C
5	A	G	A	G	T	G	A	G	T
6	C	A	G	G	T	A	G	A	G
7	A	C	T	G	T	A	C	G	T
8	C	T	G	G	T	G	A	G	T
9	T	A	T	G	T	A	A	G	T
10	C	G	G	G	T	G	A	G	C

Matrix sp

(b)

rows(sb) = 14

The background model

	1	2	3	4	5	6	7	8	9
1	A	A	A	A	A	A	C	A	G
2	A	A	A	A	A	C	A	G	G
3	A	A	A	A	C	A	G	G	T
4	A	A	A	C	A	G	G	T	G
5	A	A	C	A	G	G	T	G	A
6	A	C	A	G	G	T	G	A	G
7	C	A	G	G	T	G	A	G	T
8	A	G	G	T	G	A	G	T	A
9	G	G	T	G	A	G	T	A	A
10	G	T	G	A	G	T	A	A	A
11	T	G	A	G	T	A	A	A	A
12	G	A	G	T	A	A	A	A	A
13	A	G	T	A	A	A	A	A	A
14	G	T	A	A	A	A	A	A	A

Matrix sb

Matrix p (PFM)

	1	2	3	4	5	6	7	8	9
G	1	2	5	10	1e-17	5	1	8	1
A	3	4	2	1e-17	1e-17	5	7	1	1e-17
T	2	3	2	1e-17	10	1e-17	1e-17	1	7
C	4	1	1	1e-17	1e-17	1e-17	2	1e-17	2

Matrix b (PFM)

	1	2	3	4	5	6	7	8	9
G	4	4	4	4	4	4	4	4	4
A	8	7	7	7	7	7	7	8	8
T	1	2	2	2	2	2	2	2	2
C	1	1	1	1	1	1	1	1e-17	1e-17

$$p[i, j] = \ln \left(\frac{(p[i, j] / \text{rows}[sp])}{(b[i, j] / \text{rows}[sb])} \right)$$

(c)

Matrix p (PWM)

	1	2	3	4	5	6	7	8	9
G	-1.05	-0.36	0.56	1.25	-40.19	0.56	-1.05	1.03	-1.05
A	-0.64	-0.22	-0.92	-40.75	-40.75	0.00	0.34	-1.74	-40.89
T	1.03	0.74	0.34	-39.50	1.95	-39.50	-39.50	-0.36	1.59
C	1.72	0.34	0.34	-38.81	-38.81	-38.81	1.03	0.34	40.17

Dynamic PWMs

Dynamic PWMs. (a) Shows the motif set and the transformation of this set into a PFM stored in matrix p, (b) shows the background set and the transformation of this set into a PFM stored in matrix b. (c) It shows the construction of the PWM with the help of the two PFMs. Note that in the lower right corner of the PWM there is a value of +40.17. This value occurs due to an insufficient variation in the background set. In other words, the z sequence in our example is either too short or poor in information content.



Cite this work as:

Paul A. Gagniuc. *Algorithms in Bioinformatics: Theory and Implementation*. John Wiley & Sons, Hoboken, NJ, USA, 2021, ISBN: 9781119697961.