

Additional File 2

A

Sequence	Position														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
<i>paaA</i>	A	C	C	G	A	C	C	G	G	T	C	G	G	T	T
<i>paaH</i>	A	C	C	G	A	C	C	G	G	T	C	G	G	T	T
<i>paaZ</i>	A	C	C	A	A	C	C	G	G	T	C	G	G	T	T
<i>paaF</i>	A	C	C	A	A	C	C	G	G	T	C	G	G	T	T
<i>BCAL0211</i>	T	C	C	A	G	C	C	G	G	C	C	G	G	C	G

B

Nucleotide	Position														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	4	0	0	3	4	0	0	0	0	0	0	0	0	0	0
G	0	0	0	2	1	0	0	5	5	0	0	5	5	0	1
C	0	5	5	0	0	5	5	0	0	1	5	0	0	1	0
T	1	0	0	0	0	0	0	0	0	4	0	0	0	4	4

C

corrected
probability
calculation

$$p(b,i) = \frac{f(b,i) + \sqrt{N} \cdot p(b)}{\sqrt{N} + N}$$

PWM
conversion

$$W(b,i) = \log_2 \left(\frac{p(b,i)}{p(b)} \right)$$

D

Nucleotide	Position														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	1.86	-1.69	-1.69	1.48	1.86	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69
G	-1.69	-1.69	-1.69	0.19	-0.47	-1.69	-1.69	1.25	1.25	-1.69	-1.69	1.25	1.25	-1.69	-0.47
C	-1.69	1.25	1.25	-1.69	-1.69	1.25	1.25	-1.69	-1.69	-0.47	1.25	-1.69	-1.67	-0.47	-1.69
T	0.19	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	1.86	-1.69	-1.69	-1.69	1.86	1.86