

SUPPLEMENTAL MATERIAL 3

For the paper “Quality estimation of multiple sequence alignments by Bayesian hypothesis testing”

A. Tomovic and E. J. Oakeley

Table 3. Properties of random and JASPAR alignments

Alignments	Average length	Average number of sequences involved in the alignment
Random alignments	11	32
JASPAR alignments	10	30

The 95% posterior interval for the differences between the mean number of sequences in two groups is [-3.4, 6.8] and the 95% posterior interval for the difference between the mean length of two groups is [-0.6, 1.7]. The posterior intervals were calculated using WinBUGS (Spiegelhalter, et al., 2004). Both posterior intervals include 0 therefore there is no significant difference in the average length and the number of sequences within the two groups (random alignments and JASPAR alignments).

WinBUGS code for the calculation of the 95% posterior interval for the difference between the means of the random and JASPAR alignments

```
Model{  
  
for (i in 1:N)  
{  
  random[i]~dnorm(mu[1],tau[1])  
}  
  
for (i in 1:M)  
{  
  jasper[i]~dnorm(mu[2],tau[2])  
}  
  
for (i in 1:2)  
{  
  mu[i]~dflat()  
  tau[i]~dgamma(1,1)  
}  
  
difference<-mu[1]-mu[2]  
}
```

Input data:

N- size of vector random (number of random sequences, N=100)

M-size of vector jasper (number of jasper alignments, M=107)

random – vector of lengths (or number of sequences) for random alignments

jaspar - vector of lengths (or number of sequences) for jasper alignments