

# Randomized Info

Christian Otto

The randomized approach to motif discovery begins by randomly selecting motifs of length  $k$  in each DNA sequence. We calculate the score of this collection, then use the motifs to generate a profile matrix. We then use this profile matrix to generate a new collection of motifs from the DNA sequences. We compare this score to the previous score of motifs, and if the score is better, we keep the new collection of motifs. We then use this new collection of motifs to generate the profile-matrix and generate another new collection of motifs. We continue this process until the score does not improve and we keep this final collection of motifs as the global best motifs. We then restart the entire process with a new randomly selected collection of motifs and compare it to the global best motifs. We do this 3,000 times, ultimately keeping the best collection of motifs as our estimated motifs. (Compeau and Pevzner 2018)

## References

Compeau, Phillip, and Pavel Pevzner. 2018. *Bioinformatics Algorithms: An Active Learning Approach*. 3rd ed. Active Learning Publishers.