

Bioinformatics I & II

Wayne Delpont & Sergei Kosakovsky Pond

Fun Exercise:

Richard Dawkins in his book *The Blind Watchmaker* ponders on whether a monkey can produce the complete works of Shakespeare by randomly banging away at keys on a keyboard. He also proposes to make the problem simpler by evaluating whether a monkey can reproduce, at random, just one line from Hamlet.

”METHINKS IT IS LIKE A WEASEL”

Your computer is your monkey.

(i) Write a piece of code that generates a random sequence of letters from an alphabet which comprises 26 capital letters and a space. Score the randomly generated sequence by counting the number of matches with the target sequence above. What is the best score you can achieve with 1000 runs of the program?

(ii) Modify your code to allow for cumulative mutation, i.e. start with a random string and mutate one position per generation. Generate 100 mutated strings every generation and choose the string with the highest fitness (best score when compared against the target) to be the input for the next generation (i.e the one that will be mutated). How long (in generations) does it take to generate the target string?

This process is analogous to the evolutionary process, although in evolution there is no knowledge of the target. Although this process is a simplification it demonstrates that complex features can evolve through stochastic mutations if those organisms which carry the mutations bear some fitness advantage.